Fall 2019

Modeling the Occurrence of Terrorist Attacks

Earl Hur
Iowa State University, earl88@iastate.edu

Modeling the Occurrence of Terrorist Attacks

Earl Hur

November 30, 2019

Abstract

"From 1970 to 2015, there were 7,584 terrorist attacks in Latin America and Asia. We investigate modeling these events using dynamic statistical models with a monthly time step. Methodologically, dynamic models are the most straight forward when based entirely on normal probability structures. A potential complication is that the number of terrorist attacks are counts, with a substantial number of zero values when considered on a monthly basis. We consider a traditional additive error dynamic model in which the mean process evolves through time following an autoregressive structure. The latent process follows normal distributions with these means. The actual observation process is then a discretized version of the latent process. We contrast this model with a model that takes the observation process to follow Poisson distribution directly. The estimation and inference proceed via Markov Chain Monte Carlo methods and the models are assessed based on the combination of the autocorrelations and the maximum number of recorded attacks."

1. Introduction

The Global Terrorism Database (GTD) contains information on terrorist attacks around the world from 1970 to 2018. There exists a similar pattern in the total number of monthly attacks within Latin American countries based on these data. There also exists another similar pattern in the total number of monthly attacks within Eurasian countries based on the same dataset. Eurasian countries show a temporal trend in the total number of monthly attacks in countries from non-Latin America including Burkina Faso, Central

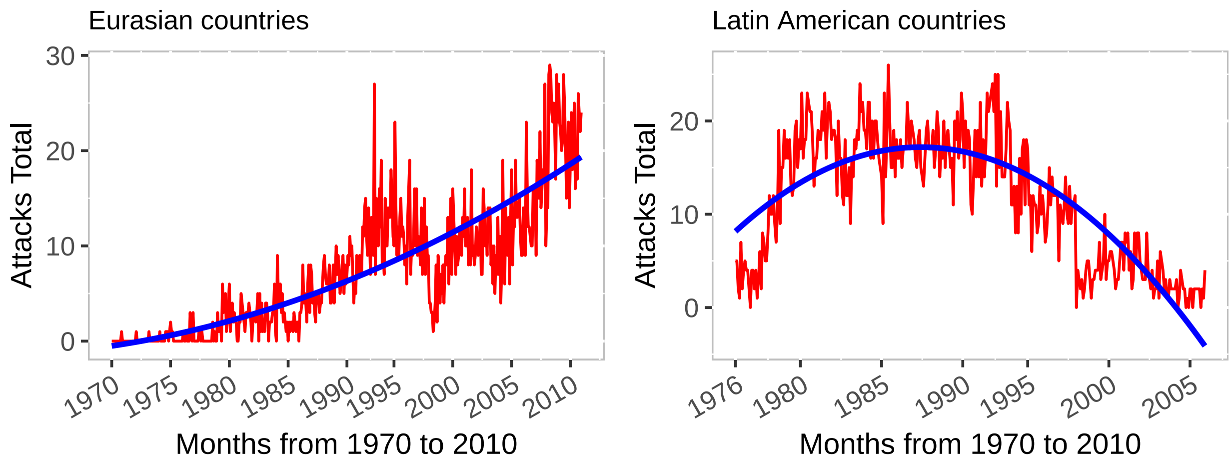


Figure 1: Quadratic fits of the two temporal patterns

African Republic, India, Iraq, Mali, Nigeria, Pakistan, Sudan, Syria, Tunisia, Ukraine, and Yemen (left side panel of Figure 1). Latin American countries show another temporal trend of the total number of monthly attacks in countries of Latin America that includes Bolivia, Brazil, Chile, Colombia, El Salvador, Guatemala, Haiti, Nicaragua, and Peru (right side panel of Figure 1). For Eurasian countries, the monthly number of attacks kept increasing from 1970 to 2010. Latin American countries show a drastic increase late in the 1970s and steadly decreases after about 1990.

This paper develops theoretical models to describe the number of monthly terrorist attacks from the countries in these two patterns. The two models are both based on the AR(1) structure. The first model considered a traditional additive error dynamic model in which the latent mean process evolves through time. The actual observed process is a discretized version of the latent process. We contrast this model with a model that takes the observed process to follow the Poisson distribution directly. Estimation and inference proceed via Markov Chain Monte Carlo methods, and the models are assessed based on a combination of the maximum number of recorded attacks and the ability of the models to adquately represent temporal structure in the data.

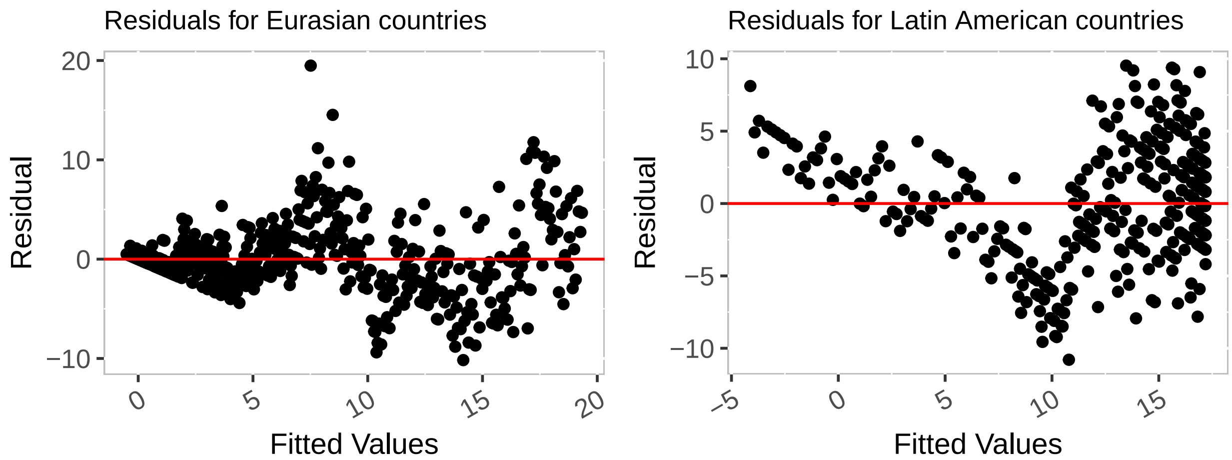


Figure 2: Residual plots for the two patterns of attacks

2. Exploratory Analyis

Linear regressions with quadratic response functions were fitted to the data from the two regions to see the overall patterns in the number of monthly terrorist attacks. The model with this quadratic fit would be:

$$Y_{(t)} = \beta_0 + \beta_1 t + \beta_2 t^2 + \epsilon_{(t)}, \qquad \epsilon_{(t)} \quad i.i.d \sim N(0, \sigma^2) \tag{1}$$

The two regression lines appear to describe the overall trends in the data (Figure 1). The first fitted line (left side of Figure 1) based on attacks in Eurasian countries is steadily increasing. The second fitted line (right side of Figure 1) based on attacks in Latin American countries has a peak in mid 1970s and then decreases until 2005. After investigating the residual plots from the regression lines with quadratic terms, we formulated several statistical models, taking into consideration the underlying structures that may remain after the regression.

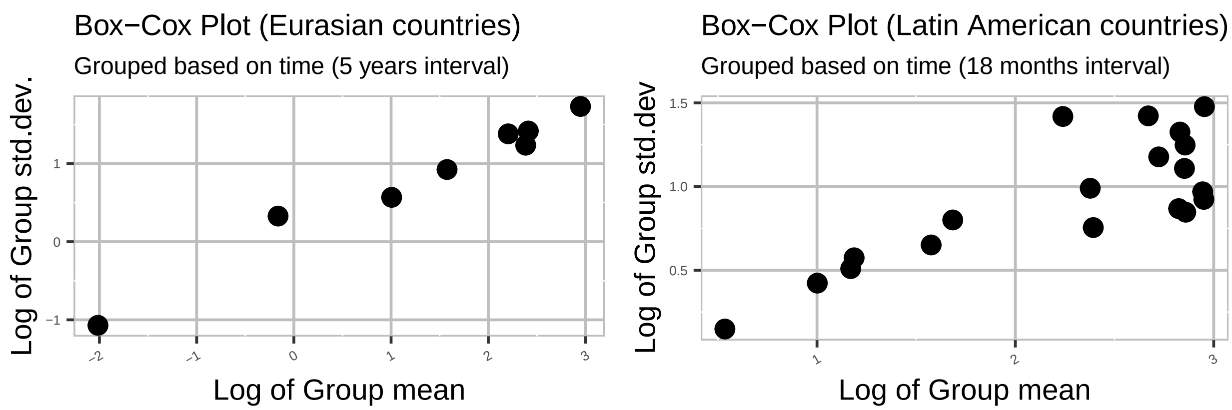


Figure 3: Box-Cox plots for the temporal patterns

2.1 Residual Plots

In this section, we summarize patterns in the residuals using graphical display, such as residual plots and Box-Cox plots. Our objective is to determine whether important patterns that are inherent in the data remain after removal of overall trend by the regressions. Plots of residuals versus fitted values for the regressions in Eurasia and Latin America are presented in Figure 2. The residuals are not randomly scattered around zero, but they rather seem to exibit non-constant variance. Although this non-homogeneous variance issue looks more obvious in Eurasian countries, data for Latin American countries also have the same issue in large fitted values.

The Box-Cox plots are shown in Figure 3 to see if the non-homogeneous variance could be modeled as a power of the means in the number of terrorist attacks. We grouped the observations, which is total number of attacks by months, into equal time intervals and computed the group means and standard deviations. By plotting the log of grouped means and standard deviations, Figure 3 suggests that there exist a mean and variance relationships in the number of terrorist attacks (the log of grouped standard deviation is proportional to that of the mean).

The least squares slope for the Box-Cox plot for Eurasian countries (left panel in Figure 1)

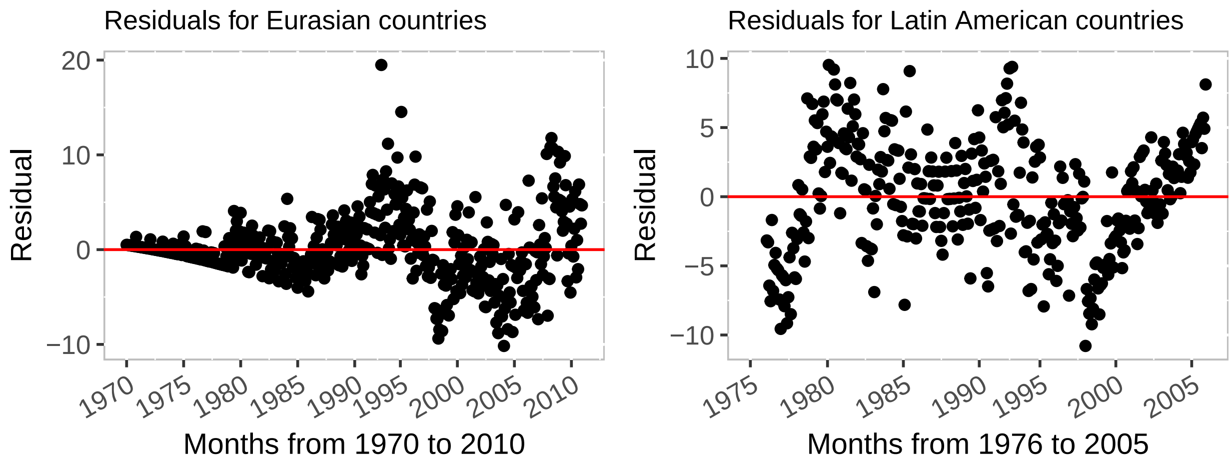


Figure 4: Residual plots for the two patterns of attacks

was 0.53. This implies that if Y_t is the number of terrorist attacks for month t in this group of countries, then the variance of Y_t can be modeled as $Var(Y_t) = [E(Y_t)]^{2\times 0.53}$. Similary, a least squares slope for the plot for Latin American countries was 0.39, so that the variances in a model for this group of countries might be represented as $Var(Y_t) = [E(Y_t)]^{2\times 0.39}$. Since the slopes are both reasonably close to 0.5, that suggests variances in both groups might be modeled as proportional to the expected values as an initial step.

2.2 Autocorrelation Function

Residuals can also be plotted against time, instead of fitted value (Figure 4). Using these residual plots from two different regions, we can see that both plots hint at some remaining temporal structure after the regression. One of the tools to find the temporal structures of the number of attacks is the autocorrelation function, as it measures the internal association between observations in time. By looking at the autocorrelation function and the partial autocorrelation function plots from the two regions (Figure 5, Eurasian countries in the upper panels and Latin American countries in the lower panels), it appears that the remaining temporal process after the quadratic fit of terrorist attacks has an autocorrelation between near-time observations. Although it apprars that dependence

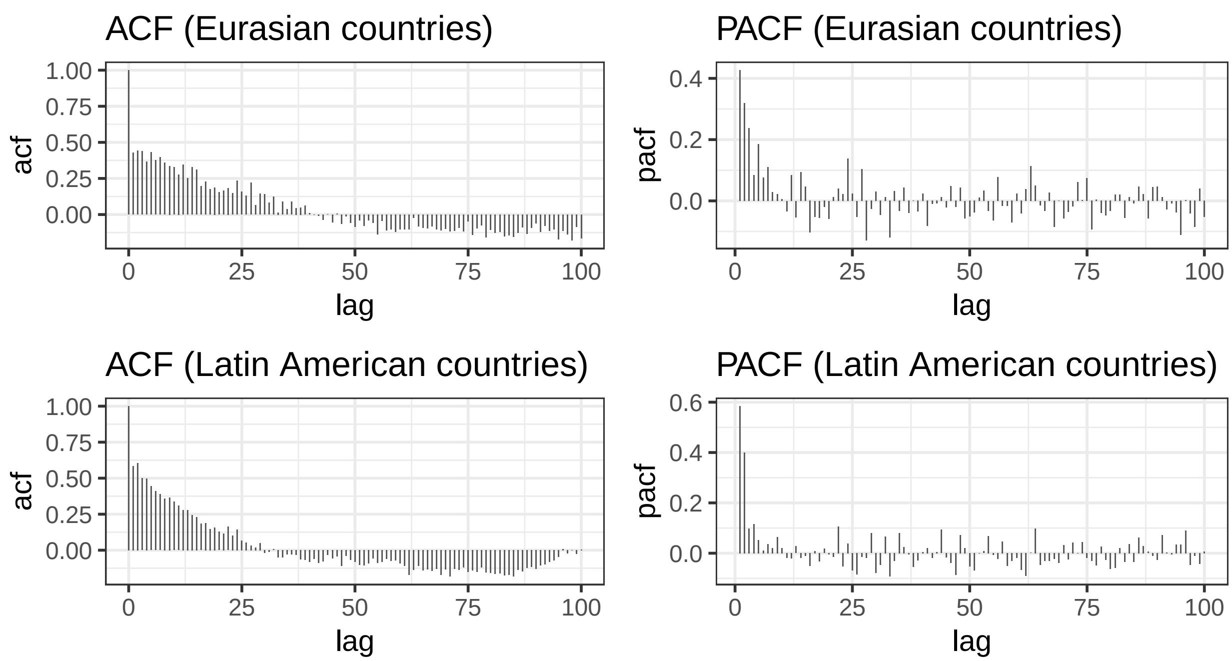


Figure 5: autocorrelation and Partial autocorrelation Function

may be more complex, for simplicity we use the first-order autocorrelation structure, AR(1), for both regions in modelling the monthly number of terrorist attacks.

Exploratory analysis has revealed two structures in residuals after the quadratic regression: variances that are proportional to the means, and autocorrelation over time. In the next section, we will formulate models that attempt to account for these two structures.

3. Model Formulation

The objective in this section is to develop models by modifying the simple linear regression with quadratic expections to incorporate nonconstant variances and temporal dependencies. We will consider these two issues in turn.

3.1 Modeling Non-Constant Variances

The Box-Cox plots of Figure 3 suggest that variances in both groups of countries might be adquately modeled as proportional to the expected values. Incorporating this into the simple quadratic regression models we have:

$$Y_{(t)} = \beta_0 + \beta_1 t + \beta_2 t^2 + |\mu|^{1/2}\epsilon_{(t)}, \qquad \epsilon_{(t)} \;\; i.i.d \sim N(0, \sigma^2) \tag{2}$$

3.2 Modeling Temporal Dependence

There are two options to incorporate the first-order temporal structure in our model. Incorporating temporal dependence into the error process of the response model results in:

$$Y_{(t)} = \mu_{(t)} + w_{(t)}, \quad \text{where} \quad w_{(t)} = \gamma w_{(t-1)} + v_{(t)}, \qquad v_{(t)} \;\; i.i.d \sim N(0, \sigma^2) \tag{3}$$

This model implies that

$$\begin{aligned} var(Y_{(t)}) &= \frac{\sigma^2}{1-\gamma^2} \\ cov(Y_{(t-1)}, Y_{(t)}) &= \frac{\gamma\sigma^2}{1-\gamma^2} \\ cov(Y_{(t-k)}, Y_{(t)}) &= \frac{\gamma^k\sigma^2}{1-\gamma^2} \end{aligned}$$

An alternative is to incorporate temporal dependence into the expectation process of the response model, which gives:

$$Y_{(t)} = \mu_{(t)} + \epsilon_{(t)}, \quad \text{where} \quad \mu_{(t)} = \lambda_{(t)} + \gamma(\mu_{(t-1)} - \lambda_{(t-1)}) + v_{(t)}, \tag{4}$$

$$\epsilon_{(t)} \;\; i.i.d \sim N(0, \sigma^2) \quad \text{and} \quad v_{(t)} \;\; i.i.d \sim N(0, \tau^2)$$

This then implies that:

$$var(Y_{(t)}) = \frac{\tau^2}{1-\gamma^2} + \sigma^2$$
$$cov(Y_{(t-1)}, Y_{(t)}) = \frac{\gamma\tau^2}{1-\gamma^2}$$
$$cov(Y_{(t-k)}, Y_{(t)}) = \frac{\gamma^k\tau^2}{1-\gamma^2}$$

Notice the fact that the variance of the observations for model (3) does not contain an extra additive term but model (4) does. This is illustrated in Figure 6, which presents the autocorrelation functions for two simulated sets of data: the left pannel is from model (3) and the right pannel is from model (4). In these simulated examples, $\mu = 0.5$ for all t, $\sigma^2 = 0.7$, and $\gamma = 0.4$. We can see that both autocorrelation function plots in Figure 5 for the actual data are similar to the plot for model (4) (right pannel) in Figure 6. Therefore, modelling the temporal structure into the expectation process will be more reasonable.

3.3 Potential Models.

In Section 3.1, we concluded that the variances might be modeled as proportional to the expected values. In Section 3.2, we showed that the temporal structure might be incorporated through the first-order autocorrelation in the expection function, similar to model (4). We would like to combine these two issues in modeling the monthly number of terrorist attacks. We will present two potential model structures to accomplish this in the next two sub-sections.

3.3.1 Gaussian Latent Variable Model.

To incorporate both non-constant variances and temporal structure, we finalize model formulation by defining the observational process. In this section, we consider using normal distributions for the observation process, while the mean process evolves through

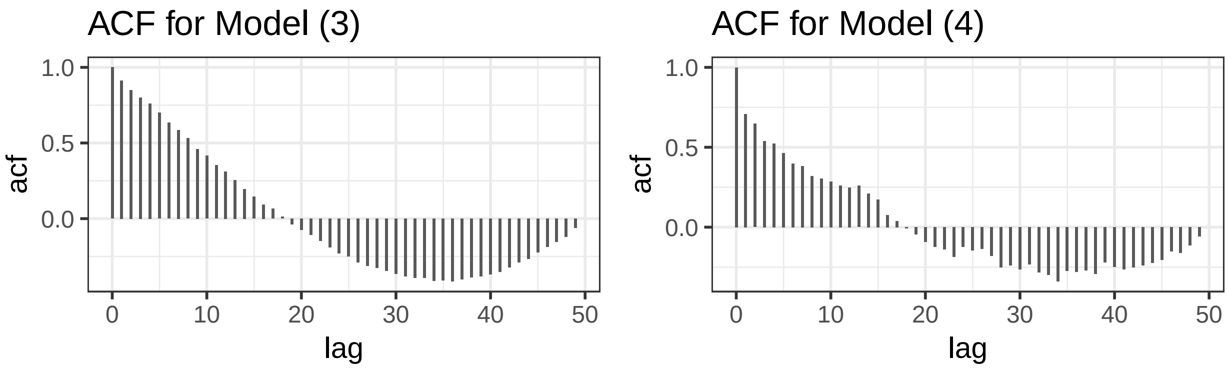


Figure 6: autocorrelation for Two Options

time following an autoregressive structure similar to model (4). The model becomes:

$$Y_{(t)} = \mu_{(t)} + \epsilon_{(t)}|\mu_{(t)}|^{1/2}\ ,\ \ \epsilon_{(t)} \sim N(0, \sigma^2) \tag{5}$$

where $\mu_{(t)} = \lambda_{(t)} + \gamma(\mu_{(t-1)} - \lambda_{(t-1)}) + v_{(t)},\ \ v_{(t)}\ \ \text{i.i.d} \sim N(0, \tau^2)$
and $\lambda_{(t)} = \beta_0 + \beta_1 t + \beta_2 t^2$

This model implies that:

$$Y_{(t)}|\mu_{(t)}\ \ \text{i.i.d} \sim N(\mu_{(t)}, \sigma^2|\mu_{(t)}|)$$
$$\mu_{(t)}|\mu_{(t-1)}\ \ \text{i.i.d} \sim N\Big(\lambda_{(t)} + \gamma(\mu_{(t-1)} - \lambda_{(t-1)}), \tau^2)\Big),$$
$$\text{and } \mu_{(t)}\ \ \text{i.i.d} \sim N\Big(\lambda_{(t)}, \frac{\tau^2}{1-\gamma^2}\Big),\quad \text{which is the marginal distribution of } \ \mu_{(t)}$$

$$Var(Y_{(t)}) = E\Big(Var(Y_{(t)}|\mu_{(t)})\Big) + Var\Big(E(Y_{(t)}|\mu_{(t)})\Big)$$
$$= Var(\mu_{(t)}) + \mu_{(t)}\sigma^2 = \frac{\tau^2}{1-\gamma^2} + \mu_{(t)}\sigma^2$$

Applying model (5) directly to the count of terrorist incidences causes an issue in that $Y_{(t)}$ in model (5) has possible values on the entire real line. However, the numbers of terrorist attacks are non-negative integer values and the data include zero values as well as other small integer values. In order to deal with this issue, we modify model (5) as follows:

$$W_{(t)} = \begin{cases} 0, \text{ if } Y_{(t)} \leq 0 \\ \lceil Y_{(t)} \rceil, \text{otherwise} \end{cases} \tag{6}$$

and we assume that $W_{(t)}$ represent the observation process. The $Y_{(t)}$ given in (5) now represents a latent process we think of as an "unobserved force of terrorism." If this unobserved force of terrorism is negative, the terrorist organization would not want to attack at all and the monthly number of terrorist attacks will become zero. This unobserved force would be realized as the number of attacks in a month.

3.3.2 Poisson Response Dynamic Model.

Now, we introduce a simpler model to compare with the model above. We use the Poisson distribution directly, which accounts for both non-negative integer valued observation process and variances that are proportional to means. This model is,

$$Y_{(t)}|\mu_{(t)} \sim Pois(\mu_{(t)}) \tag{7}$$

where $log(\mu_{(t)}) = \lambda_{(t)} + \gamma\big(log(\mu_{(t-1)}) - \lambda_{(t-1)}\big) + v_{(t)},\ \ v_{(t)} \sim N(0, \tau^2)$
and $\lambda_{(t)} = \beta_0 + \beta_1 t + \beta_2 t^2$

This model now implies that:

$$log(\mu_{(t)}) \sim N\Big(\lambda_{(t)}, \frac{\tau^2}{1-\gamma^2}\Big), \text{ which is the marginal distribution of } log(\mu_{(t)}),$$

$$\text{and } log(\mu_{(t)}|\mu_{(t-1)}) \sim N\Big(\lambda_{(t)} + \gamma(\mu_{(t-1)} - \lambda_{(t-1)}), \tau^2)\Big)$$

$$Var(Y_{(t)}) = E\Big(Var(Y_{(t)}|\mu_{(t)})\Big) + Var\Big(E(Y_{(t)}|\mu_{(t)})\Big)$$
$$= Var(\mu_{(t)}) + E(\mu_{(t)}) = e^{\left(2\lambda_{(t)}+\frac{\tau^2}{1-\gamma^2}\right)}\Big[e^{\left(\frac{\tau^2}{1-\gamma^2}\right)} - 1\Big] + e^{\left(\lambda_{(t)}+\frac{\tau^2}{2(1-\gamma^2)}\right)}$$

4. Estimation

We use Markov Chain Monte Carlo to approximate posterior distributions of the parameters of the two models for both Eurasian and Latin American countries. We use random walk Matropolis-Hastings within Gibbs algorithm for MCMC after we derive the full conditional distribution of each parameter in the two models.

4.1 Prior Distributions.

The regression parameters, β_0, β_1, β_2, both models (6) and (7) determine the the marginal expectation function $\lambda_{(t)}$. Therefore, it is natural to assign each of these parameters normal priors.

$$\beta_0 \sim N(m_0, v_0)$$

$$\beta_1 \sim N(m_1, v_1)$$

$$\beta_2 \sim N(m_2, v_2)$$

It is common to set m_0, m_1, and m_2 to 0 and take large values for the variances v_0, v_1, and v_2, which we set to 25.

For the autoregressive parameter γ, a uniform prior was chosen over the parameter space of $(-1, 1)$ as we do not have any information about what the correlation between the near time values in mean process.

$$\gamma \sim Unif(-1, 1)$$

The variance of the expectation process $\mu_{(t)}$is given by the parameter τ^2. Following the suggestions of Gelman (2006), we used the uniform prior on τ. The upper bound of this uniform distribution is arbitrary but may be assessed via sensitivity analysis.

$$\tau \sim Unif(0, u)$$

The prior for σ^2 is chosen as inverse gamma distribution based on conditional conjugacy in the variance of the observation process.

$$\sigma^2 \sim inverse.gamma(a, b)$$

where the values $a = 1, b = 1$ were used.

For the Poisson response dynamic model, we used the same prior distributions with no σ^2. Using these prior distributions, the full conditional posterior distributions are given in the next section.

4.2 Full Conditional Posterior Distributions

For the Gaussian latent variable model, the procedure for deriving the full conditional posterior distributions for each parameter is given in Appendix A1. The full conditional posterior distributions for Model (6) are:

$$p(\beta_0|\cdot) \propto \pi(\beta_0) \times g(\underset{\sim}{\mu}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2) \times g(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2)$$

$$p(\beta_1|\cdot) \propto \pi(\beta_1) \times g(\underset{\sim}{\mu}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2)$$

$$p(\beta_2|\cdot) \propto \pi(\beta_2) \times g(\underset{\sim}{\mu}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2)$$

$$p(\gamma|\cdot) \propto \pi(\gamma) \times g(\underset{\sim}{\mu}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2) \times g(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2)$$

$$p(\tau^2|\cdot) \propto \pi(\tau^2) \times g(\underset{\sim}{\mu}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2) \times g(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, \tau^2)$$

$$p(\sigma^2|\cdot) \propto \pi(\sigma^2) \times f(\underset{\sim}{y}|\underset{\sim}{\mu}, \sigma^2)$$

$$p(\mu_{(0)}|\cdot) \propto g(\mu_{(0)}) \times g(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, \tau^2) \times h(y_{(0)}|\mu_{(0)}, \sigma^2)$$

$$p(y_{(0)}|\cdot) \propto h(y_{(0)}|\mu_{(0)}, \sigma^2)$$

Finally, for $t = 1, 2, 3, ..., T$,

$$p(\mu_{(t)}|\cdot) \propto g(\mu_{(t)}|\mu_{(t-1)}) \times g(\mu_{(t+1)}|\mu_{(t)}) \times h(y_{(t)}|\mu_{(t)}, \sigma^2)$$

$$p(y_{(t)}|\cdot) \propto h(y_{(t)}|\mu_{(t)}, \sigma^2) \times q(w_{(t)}|y_{(t)})$$

The parameters with known full conditional distributions are:

$$\beta_0|\cdot \sim N(M_0, S_0^2)$$

$$\beta_1|\cdot \sim N(M_1, S_1^2)$$

$$\beta_2|\cdot \sim N(M_2, S_2^2)$$

$$\tau^2 = w|\cdot \sim inverse.gamma(A_1, B_1)$$

$$\sigma^2|\cdot \sim inverse.gamma(A_2, B_2)$$

$$y_{(0)}|\cdot \sim N(M_3, S_3^2)$$

In the GIbbs algorithm, we directly sample from the above full conditional posterior distributions for model (6). The remaining full conditional posteriors for $\gamma, \mu_{(0)}, \mu_{(t)},$ and $Y_{(t)}$ were sampled from Metropolis-Hastings. For the Poisson response dymanic model, the full conditional posterior distributions are given in Appendix A2. The parameters with known full conditional distributions are:

$$\beta_0|\cdot \sim N(M_4, S_4^2)$$

$$\beta_1|\cdot \sim N(M_5, S_5^2)$$

$$\beta_2|\cdot \sim N(M_6, S_6^2)$$

$$\tau^2|\cdot \sim inverse.gamma(A_3, B_3)$$

$$\mu_{(0)} \sim N(M_7, S_7^2)$$

For the above parameters, we use the Gibbs algorithm directly to sample from the full conditional posterior distributions for model (7). The parameters for above distributions are given in Appendix A.2 as well. The remaining full conditional posteriors for $\gamma, \mu_{(0)}, \mu_{(t)}$, and $y_{(t)}$ were sampled from Metropolos-Hastings using random walk jump proposals.

4.3 Markov Chain Monte Carlo Method

For each model, we sampled 20,000 values with the burn-in of 2,000 from the joint posterior distributions using the algorithm described previously. In the following sections, we will discuss what we found from the posterior distributions for each parameter.

4.3.1 Details of simulation procedure

The trace plots and autocorrelation function plots for β_1, γ and σ^2 for the Gaussian latent variable model (6) were presented in Figure 7. For the Poisson response dynamic model (7), the trace plots and autocorrelation plots for the same parameters are presented in Figure 8. The trace plots and autocorrelation plots for the other parameters are shown in Appendix C. Based on Figure 7 and Figure 8, the slow mixing for the Poisson response dynamic model may temper any strong conclusions about the posterior distribution.

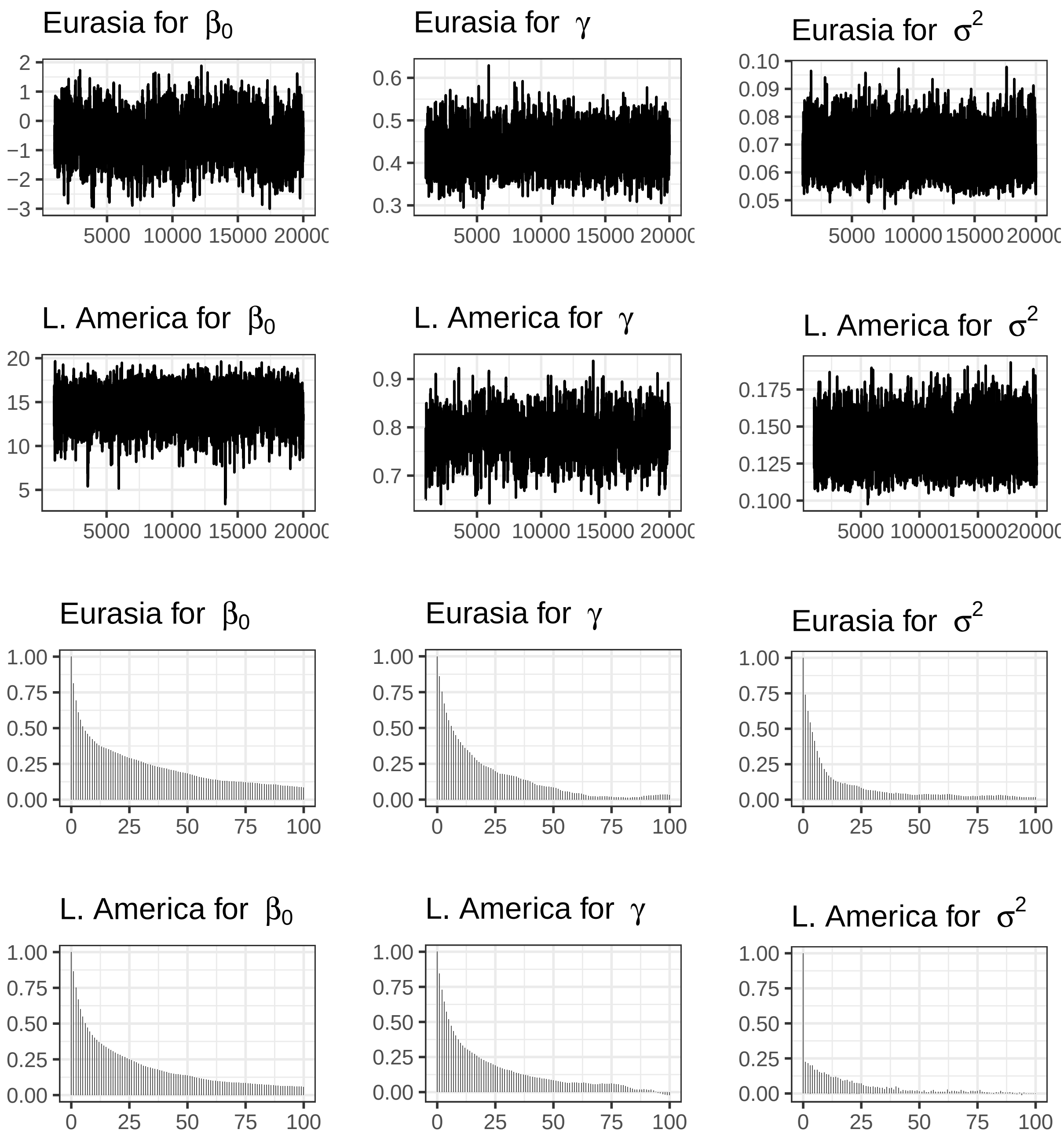


Figure 7: Trace and Autocorrelation Plots for Model (6)

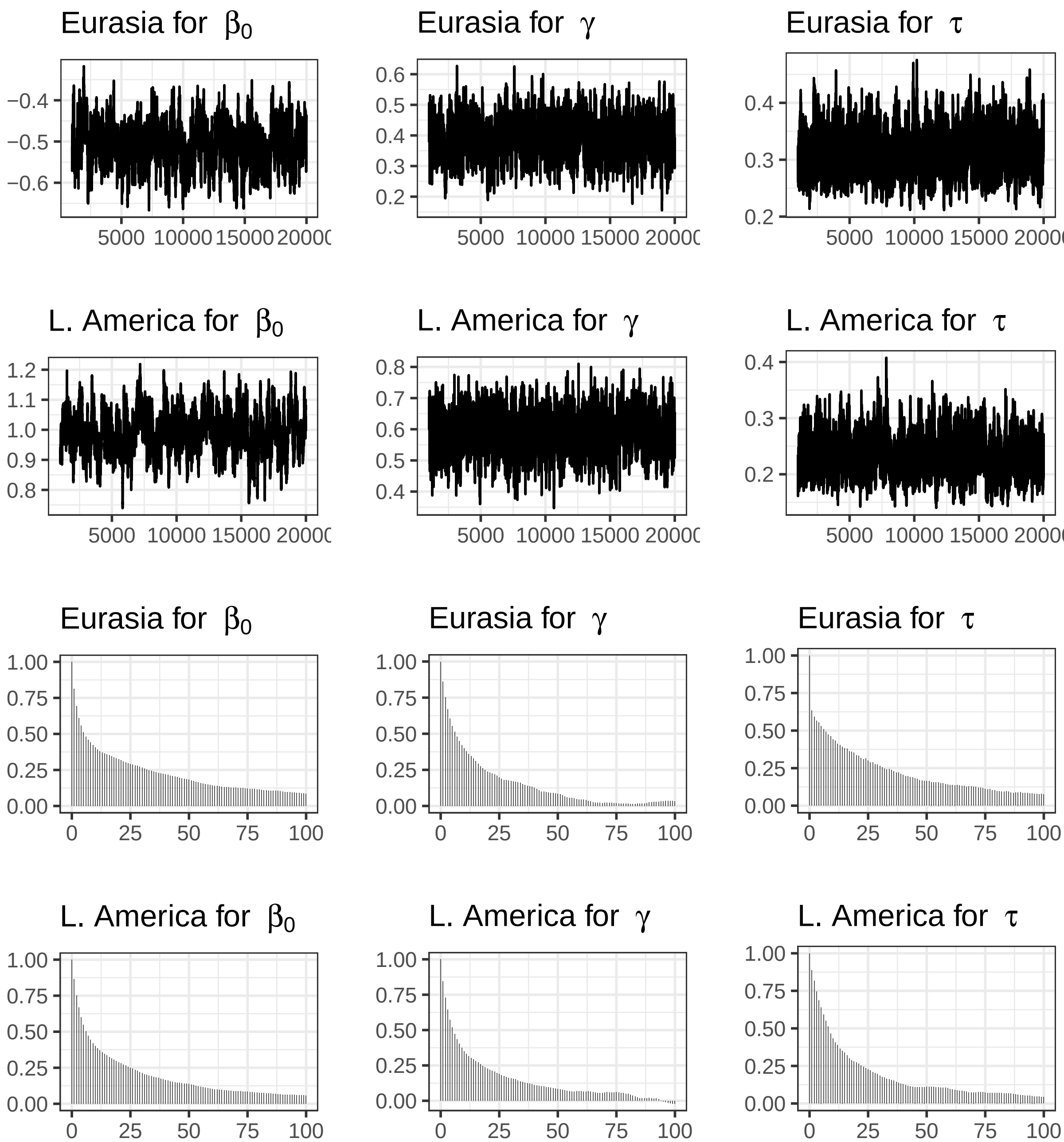


Figure 8: Trace and Autocorrelation Plots for Model (7)

Table 1: Posterior Summaries for Model (6) fit to Eurasia

Model (6), Eurasia	minimum	Q1	Q2	Q3	Maximun	mean
β_0	-3.001	-1.002	-0.562	-0.120	1.883	-0.561
β_1	-3.098	6.591	8.347	10.082	18.512	8.326
β_2	0.994	9.625	11.315	13.035	22.287	11.334
γ	0.292	0.406	0.434	0.463	0.629	0.435
τ^2	8.517	10.403	10.860	11.349	14.424	10.893
σ^2	0.047	0.063	0.067	0.072	0.098	0.068

Table 2: Posterior Summaries for Model (6) fit to Latin America

Model (6), L. America	minimum	Q1	Q2	Q3	Maximun	mean
β_0	3.369	13.460	14.494	15.496	19.635	14.418
β_1	-11.721	-1.191	1.411	4.004	13.653	1.378
β_2	-25.558	-13.569	-10.718	-7.704	10.329	-10.555
γ	0.641	0.757	0.784	0.813	0.938	0.784
τ^2	9.341	11.634	12.261	12.933	17.169	12.313
σ^2	0.098	0.131	0.138	0.146	0.193	0.139

4.3.2 Monte Carlo Results

The results from the Markov Chain Monte Carlo (MCMC) simulation from the Gaussian latent variable model for Eurasian countries are presented in Table 1. For Latin American countries, the MCMC results are shown in Table 2.

The regression parameters, β_0, β_1, and β_2, in Tables 1 and 2 determine the pattern of the mean process in the number of monthly terrorist attacks. Because polynomials were used as a flexible device to describe overall trend, individual regression coefficient estimates do not have any particular interpretation or meaning. By looking at the posterior mean values for these parameters, the mean process in monthly number of terrorist attacks is increasing when time, t, is increasing. For Latin American countries, the pattern goes up and down. The autoregressive coefficient for Latin American countries is higher than that of Eurasian countries. That being said, the means of near-time monthly number of terrorist attacks are more correlated in Latin American countries than in Eurasian countries. The variance

Table 3: Posterior Summaries for Model (7) fit to Eurasia

Model (7), Eurasia	minimum	Q1	Q2	Q3	Maximun	mean
β_0	-0.667	-0.543	-0.507	-0.473	-0.351	-0.508
β_1	2.338	2.813	2.946	3.092	3.509	2.950
β_2	-0.507	-0.149	-0.045	0.053	0.382	-0.045
γ	0.155	0.350	0.393	0.434	0.627	0.392
τ^2	0.211	0.288	0.310	0.333	0.476	0.311

Table 4: Posterior Summaries for Model (7) fit to Latin America

Model (7), L. America	minimum	Q1	Q2	Q3	Maximun	mean
β_0	0.739	0.947	0.991	1.038	1.219	0.992
β_1	4.322	4.969	5.134	5.308	5.795	5.131
β_2	-7.199	-6.647	-6.470	-6.285	-5.473	-6.459
γ	0.347	0.548	0.589	0.631	0.810	0.589
τ^2	0.140	0.209	0.229	0.250	0.408	0.231

of the mean process in monthly number of attacks, however, is not dramatically different between the two from Tables 1 and 2. By comparing values of σ^2, the latent process seems to account for more of the small-scale structure in the data from Eurasia than in Latin America. Additionally, the overall data structure seems to be captured more entirely by the polynomial fit in the case of Eurasia as compared to Latin America.

The results from the MCMC simulation for the Poisson random variable model from Eurasian countries are presented in Table 3. For Latin American countries, the MCMC results are shown in Table 4.

Again for the Poisson response variable model, the autoregressive nature of the mean process for Eurasia is not strong compared to that of Latin American countries. On the other hand, the variance of the mean process in Model (6) is greater than that in Model (7). Model (7) places a constraint on the variance because of using the Poisson response variable which is more restrictive compared to the model with Gaussian latent variables.

Table 5: Variances of each parameter value

Model (6), Eurasia	NoThin	Thin24	Thin48	Thin96
β_0	0.427	0.439	0.44	0.416
β_1	7.087	7.186	7.112	7.044
β_2	6.645	7.005	6.95	6.982
γ	0.002	0.002	0.002	0.002
τ^2	0.497	0.536	0.491	0.447
σ^2	<0.001	<0.001	<0.001	<0.001

5. Model Comparison

Model assessment was based on posterior predictive p-values. For exploration of posterior distributions, dependence structure in MCMC chains is not a major concern. However, posterior predictive p-values are in the form of Monte Carlo averages. Therefore we would like to have our posterior predictive datasets be independent, or at least not strongly dependent. The potentially harmful effect of dependence occurs in the computation of Monte Carlo variances.

For the Gaussian latent variable model from Eurasian countries, we took every 24th, 48th, and 96th values and calculated the variances of the Markov Chain values for each parameter with these thinned values. Table 5 presents the results. The variances presented in Table 5 are roughly the same for all thinning values. We decided to use thining of the 24th values to get the results of the Monte Carlo method and calculate the posterior predictive p-values for both models.

In this section, we use posterior predictive p-values to compare the two models in different aspects. There are various critera that we can use, such as the maximum, IQR, or autocorrelation coefficients. Each criterion assesses a different aspect of the models. If we use the maximum, then we assess whether the model captures the extreme cases of the number of monthly terrorist attacks. Using IQR, we asseesses the variability of the models. The first-, second- and third-order autocorrelation coefficients will access if the

Table 6: Posterior Predictive p-values

Criteria	Model(6), Eurasia	Model(7), Eurasia	Model(6), Latin America	Model(7), Latin America
Max. Value	0.150	0.072	0.063	0.134
IQR	0.216	0.000	0.008	0.000
1st ACF	0.361	0.024	0.039	0.019
2nd ACF	0.009	0.001	0.231	0.001
3rd ACF	0.000	0.000	0.181	0.000

models capture the time dependent structure of the terrorist attacks. Table 6 shows the posterior predictive p-values for each of these criteria. Since the p-value was calculated as the minimum proportion of test statistics from posterior predictive distribution that are smaller or larger than from the data, a small p-value means the models did not entirely capture the specified criteria from the number of monthly terrorist attacks.

Table 6 indicates that Model (6) can be used to capture the maximum values and variability in the monthly number of terrorist attacks in Eurasian countries. For Latin American countries, a Poisson response dynamic model can be used to pick up the maximun number of monthly terrorist attacks. First-order autocorrelation functions from Eurasian countries are captured from Model (6). This model captures the second- and third-order autocorrelation of the monthly number of terrorist attacks in Latin American countries. Even though Model (6) and (7) incorporated the first-order autocorrelation, the small p-values in the second- and third-order autocorrelation for Eurasian countries indicate that the models are not capturing all the temporal structure with only the first-order autocorrelations.

6. Discussion

In finding models to fit the data for the monthly number of terrorist attacks from the two regions, Eurasia and Latin America, we found there are two issues. The first issue was the non-constant variance and the second was the existance of the temporal dependence structure. In the Poisson response variable model, we dealt with these issues directly with the response distributtion. On the other hand, we modified the latent variable in hierarchical structures to deal with the issues in the Gaussian latent variable model. The Gaussian latent variable model was more flexible. That is, the latent process could be considered as unobserved force to attack by terrorist organizations. This unobserved process could be extended to the variability of the target group, such as the general public or the government, or political conditions of unrest that could attract terrorist organizations. The Gaussian latent variable model also described data better compared to the others. Moreover, the MCMC simulation from the Gaussian latent variable model mixed better than the Poisson response variable model.

Data for the number of monthly terrorist attacks contain a temporal structure. Using the Gaussian latent variable model with the first-order autocorrelation incorporated, the temporal structure was captured for Latin American countries. However, for Eurasian countries, the temporal structure was not captured well using just the first-order autocorrelation in the Gaussian latent variable model. This indicates that the temporal structure is not simple enough that the terrorist attacks are related only to the last month's attacks in Eurasian countries. We have already seen that the temporal process has more than the first-order autocorrelation from the patial autocorrelation plot for Eurasian countries. Our analysis using the Gaussian latent variable model and the Poisson response dynamic model suggests that there is a longer temporal memory connected to the terrorist attacks in Eurasian countries.

References

Gelman, Andrew. 2006. "Prior Distributions for Variance Parameters in Hierarchical Models (Comment on Article by Browne and Draper)." *Bayesian Analysis* 1 (3). International Society for Bayesian Analysis: 515–34.

Appendix

A. Deriving full conditional posterior distributions

A1. Gaussian latent variable model.

For β_0.

$$p(\beta_0|\cdot) \propto p(\beta_0) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\beta_0) \propto exp\Big(-\frac{(\beta_0 - m_0)^2}{2v_0}\Big) \propto exp\Big(-\frac{\beta_0^2 - 2m_0\beta_0}{2v_0}\Big) \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$
$$\propto exp\Big(-\frac{T(\gamma-1)^2\beta_0^2 + 2\sum_{t=1}^{T}\Big((\gamma-1)(\mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \beta_2(t-1)^2)\Big)\beta_0}{2w}\Big) \tag{2}$$

$$p(\mu_{(0)}|\beta_0, \gamma, w) \propto exp\Big(-\frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big) \propto exp\Big(-\frac{(1-\gamma^2)(\beta_0^2 - 2\mu_{(0)}\beta_0)}{2w}\Big) \tag{3}$$

By multiplying (1), (2) and (3),

$$p(\beta_0|\cdot) \propto p(\beta_0) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$
$$exp\Big(-\frac{\beta_0^2 - 2m_0\beta_0}{2v_0}\Big) \times exp\Big(-\frac{(1-\gamma^2)(\beta_0^2 - 2\mu_{(0)}\beta_0)}{2w}\Big)$$
$$\times exp\Big(-\frac{T(\gamma-1)^2\beta_0^2 + 2\sum_{t=1}^{T}\Big((\gamma-1)(\mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2)\Big)\beta_0}{2w}\Big)$$
$$= exp\Big(-\frac{\Big(w + v_0T(\gamma-1)^2 + v_0(1-\gamma^2)\Big)\beta_0^2 + 2\Big(v_0\sum_{t=1}^{T}((\gamma-1)C) - v_0(1-\gamma^2)\mu_{(0)} - m_0w\Big)\beta_0)}{2v_0w}\Big)$$

where $C = \mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2$

Now, let $w + v_0T(\gamma-1)^2 + v_0(1-\gamma^2)$ be A,

$2\Big(v_0\sum_{t=1}^{T}((\gamma-1)C) - v_0(1-\gamma^2)\mu_{(0)} - m_0w\Big)$ be B

and $2v_0w$ be D.

Then the full conditional posterior distribution for β_0 will be:

$$p(\beta_0|\cdot) \propto exp\Big(-\frac{(\beta_0 + \frac{B}{2A})^2}{2 \cdot \frac{D}{2A}}\Big)$$
$$\beta_0|\cdot \sim N\Big(-\frac{B}{2A}, \frac{D}{2A}\Big)$$

For β_1,

$$p(\beta_1|\cdot) \propto p(\beta_1) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\beta_1) \propto exp\Big(-\frac{(\beta_1 - m_1)^2}{2v_1}\Big) \propto exp\Big(-\frac{\beta_1^2 - 2m_1\beta_1}{2v_1}\Big) \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$
$$= \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \lambda_{(t)} - \gamma(\mu_{(t-1)} - \lambda_{(t-1)})\big)^2}{2w}\Big)$$
$$= exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \beta_0 - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2\big)^2}{2w}\Big)$$
$$= exp\Big(-\frac{\sum_{t=1}^{T}\Big(\big(\gamma(t-1) - t\big)\beta_1 + \big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)\Big)^2}{2w}\Big) \propto$$
$$exp\Big(-\frac{\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)\big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma(\mu_{(t-1)} - \beta_0 - \beta_2(t-1)^2)\big)\beta_1\big)}{2w}\Big) \tag{2}$$

By multiplying (1) and (2),

$$p(\beta_1|\cdot) \propto p(\beta_1) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$
$$exp\Big(-\frac{\beta_1^2 - 2m_1\beta_1}{2v_1}\Big)\times$$
$$exp\Big(-\frac{\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)\big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)\beta_1\big)}{2w}\Big)$$
$$= exp\Big(-\frac{w\beta_1^2 - 2m_1 w\beta_1 + v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C\beta_1\big)}{2v_1 w}\Big)$$
$$= exp\Big(-\frac{\big(v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2 + w\big)\beta_1^2 + 2(v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C - m_1 w)\beta_1\big)}{2v_1 w}\Big)$$

where $C = \big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)$.

Now, let $v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2 + w$ be A',

$2(v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C - m_1 w)$ be B'

and $2v_1 w$ be D.

Then the full conditional posterior distribution for β_1 will be:

$$p(\beta_1|\cdot) \propto exp\Big(-\frac{(\beta_1 + \frac{B'}{2A'})^2}{2\cdot\frac{D}{2A'}}\Big)$$
$$\beta_1|\cdot \sim N\Big(-\frac{B'}{2A'}, \frac{D}{2A'}\Big)$$

For β_2,

$$p(\beta_2|\cdot) \propto p(\beta_2) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\beta_2) \propto exp\Big(-\frac{(\beta_2 - m_2)^2}{2v_2}\Big) \propto exp\Big(-\frac{\beta_2^2 - 2m_2\beta_2}{2v_2}\Big) \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$= \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)\big)^2}{2w}\Big) =$$

$$exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \beta_0 - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2\big)^2}{2w}\Big) =$$

$$exp\Big(-\frac{\sum_{t=1}^{T}\big((\gamma(t-1)^2 - t^2)\beta_2 + \big(\mu_{(t)} - \beta_0 - \beta_1 t - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1)\big)\big)^2}{2w}\Big) =$$

$$exp\Big(-\frac{\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\beta_2^2 + 2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\big(\mu_{(t)} - \beta_0 - \beta_1 t - \gamma(\mu_{(t-1)} - \beta_0 - \beta_1(t-1))\big)\big)\beta_2}{2w}\Big) \tag{2}$$

By multiplying (1) and (2),

$$\begin{aligned}
&p(\beta_2|\cdot) \propto p(\beta_2) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
&exp\Big(-\frac{\beta_2^2 - 2m_2\beta_2}{2v_2}\Big) \times \\
&exp\Big(-\frac{\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\beta_2^2 + 2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C)\beta_2}{2w}\Big) \\
&= exp\Big(-\frac{w\beta_2^2 - 2m_2 w\beta_2 + v_2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\beta_2^2 + 2v_2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C)\beta_2}{2v_2 w}\Big) \\
&= exp\Big(-\frac{\Big(w + v_2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\Big)\beta_2^2 + 2\Big(v_2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C) - m_2 w\Big)\beta_2}{2v_2 w}\Big)
\end{aligned}$$

where $C = \big(\mu_{(t)} - \beta_0 - \beta_1 t - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1)\big)$

Now, let $w + v_2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2$ be A'',

$2\Big(v_2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C) - m_2 w\Big)$ be B''

and $2v_2 w$ be D.

Then the full conditional posterior distribution for β_2 will be:

$$\begin{aligned}
&p(\beta_2|\cdot) \propto exp\Big(-\frac{(\beta_2 + \frac{B''}{2A''})^2}{2\cdot\frac{D}{2A''}}\Big) \\
&\beta_2|\cdot \sim N\Big(-\frac{B''}{2A''}, \frac{D}{2A''}\Big)
\end{aligned}$$

For γ,

$$p(\gamma|\cdot) \propto p(\gamma) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\gamma) \propto 1 \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$= \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) \propto exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)})\big)^2}{2w}\Big)$$

$$\propto exp\Big(- \frac{\sum_{t=1}^{T} \big(\gamma(\lambda_{(t-1)} - \mu_{(t-1)}) + (\mu_{(t)} - \lambda_{(t)})\big)^2}{2w}\Big) \tag{2}$$

$$p(\mu_{(0)}|\beta_0, \gamma, w) \propto \sqrt{1-\gamma^2} exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big) \tag{3}$$

By multiplying (1), (2) and (3),

$$p(\gamma|\cdot) \propto p(\gamma) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$= \sqrt{1-\gamma^2} exp\Big(- \frac{\sum_{t=1}^{T} \big(\gamma^2(\lambda_{(t-1)} - \mu_{(t-1)})^2 + 2\gamma(\lambda_{(t-1)} - \mu_{(t-1)})(\mu_{(t)} - \lambda_{(t)})\big) - \gamma^2(\mu_{(0)}^2 - 2\mu_{(0)}\beta_0 + \beta_0^2)}{2w}\Big)$$

For $\tau^2 = w$,

$$\tau \sim Unif(0, u)$$

$$f_\tau(\tau) = \frac{1}{u}$$

$$w = \tau^2$$

$$\tau = w^{\frac{1}{2}} = g^{-1}(w)$$

$$f_w(w) = f_\tau(g^{-1}(w))|\frac{dg^{-1}(w)}{dw}| = \frac{1}{2u} w^{-\frac{1}{2}}$$

Therefore, the full conditional posterior distribution for w is:

$$p(w|\cdot) \propto p(w) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(w) \propto w^{-1/2} \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) =$$

$$\Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) \propto \big((w)^{-T/2}\big) exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2}{2w}\Big) \tag{2}$$

$$p(\mu_{(0)}|\beta_0, \gamma, w) \propto (w)^{-1/2} exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big) \tag{3}$$

By multiplying (1), (2) and (3),

$$
\begin{aligned}
&p(w|\cdot) \propto p(w) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
&((w)^{-T/2})exp\Big(- \frac{\sum_{t=1}^{T} \left(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\right)^2}{2w}\Big) \times (w)^{-1}exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big) \\
&= \left((w)^{-(T/2)-1}\right)exp\Big(- \frac{\sum_{t=1}^{T} \left(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\right)^2 + (1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big)
\end{aligned}
$$

Now, let $\sum_{t=1}^{T} \left(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\right)^2 + (1-\gamma^2)(\mu_{(0)} - \beta_0)^2$ be A'''.

Then the full conditional posterior distribution for w will be:

$$
\begin{aligned}
&p(w|\cdot) \propto \left((w)^{-(T/2)-1}\right)exp\Big(- \frac{A'''/2}{w}\Big) \\
&w = \tau^2|\cdot \sim inv.gamma\Big(\frac{T}{2}, \frac{A'''}{2}\Big)
\end{aligned}
$$

For σ^2,

$$
\begin{aligned}
p(\sigma^2|\cdot) \propto & p(\sigma^2) \times p(\tilde{y}_{(1)}, \tilde{y}_{(2)}, ..., \tilde{y}_{(T)}|\mu_{(1)}, \mu_{(2)}, ..., \mu_{(T)}, \sigma^2) \\
& p(\sigma^2) \propto (\sigma^2)^{-a-1}exp\big(- \frac{b}{\sigma^2}\big) \qquad (1) \\
& p(\tilde{y}_{(1)}, \tilde{y}_{(2)}, ..., \tilde{y}_{(T)}|\mu_{(1)}, \mu_{(2)}, ..., \mu_{(T)}, \sigma^2) \\
& = \Pi_{t=1}^{T} p(\tilde{y}_{(t)}|\mu_{(t)}, \sigma^2) \propto (\sigma^2)^{-T/2}exp\Big(- \frac{\sum_{t=1}^{T} \left(\tilde{y}_{(t)} - \mu_{(t)}\right)^2/2|\mu_{(t)}|}{\sigma^2}\Big) \qquad (2)
\end{aligned}
$$

By multiplying (1) and (2),

$$
\begin{aligned}
&p(\sigma^2|\cdot) \propto p(\sigma^2) \times p(\tilde{y}_{(1)}, \tilde{y}_{(2)}, ..., \tilde{y}_{(T)}|\mu_{(1)}, \mu_{(2)}, ..., \mu_{(T)}, \sigma^2) \\
&= (\sigma^2)^{-(T/2)-a-1}exp\Big(- \frac{\sum_{t=1}^{T} \Big((\tilde{y}_{(t)} - \mu_{(t)})^2/2|\mu_{(t)}|\Big) + b}{\sigma^2}\Big)
\end{aligned}
$$

Now, let $\sum_{t=1}^{T} \left(y_{(t)} - \mu_{(t)}\right)^2/2\mu_{(t)} + b$ be A''''.

Then the full conditional posterior distribution for σ^2 will be:

$$
\begin{aligned}
&p(\sigma^2|\cdot) \propto \left((\sigma^2)^{-(T/2)-a-1}\right)exp\Big(- \frac{A''''}{\sigma^2}\Big) \\
&\sigma^2|\cdot \sim inv.gamma\Big(T/2 + a, A''''\Big)
\end{aligned}
$$

For $\mu_{(0)}$,

$$p(\mu_{(0)}|\cdot) \propto p(\mu_{(0)}) \times p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w) \times p(\tilde{y}_{(0)}|\mu_{(0)})$$

$$p(\mu_{(0)}) \propto exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \lambda_{(0)})^2}{2w}\Big) \propto exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)}^2 - 2\lambda_{(0)}\mu_{(0)})}{2w}\Big) \quad (1)$$

$$p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w) \propto exp\Big(- \frac{\big(\mu_{(1)} - \lambda_{(1)} - \gamma\mu_{(0)} + \gamma\lambda_{(0)}\big)^2}{2w}\Big)$$

$$\propto exp\Big(- \frac{\gamma^2\mu_{(0)}^2 - 2\gamma(\mu_{(1)} - \lambda_{(1)} + \gamma\lambda_{(0)})\mu_{(0)}}{2w}\Big) \quad (2)$$

$$p(\tilde{y}_{(0)}|\mu_{(0)}) \propto \frac{1}{\sqrt{|\mu_{(0)}|}} exp\Big(- \frac{(\tilde{y}_{(0)}^2 - 2\tilde{y}_{(0)}\mu_{(0)})}{2\sigma^2|\mu_{(0)}|}\Big) \quad (3)$$

By multiplying (1) , (2) and (3),

$$p(\mu_{(0)}|\cdot) \propto p(\mu_{(0)}) \times p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w) \times p(\tilde{y}_{(0)}|\mu_{(0)})$$

$$\propto \frac{1}{\sqrt{|\mu_{(0)}|}} exp\Big(- \frac{\mu_{(0)}^2 + 2(\gamma\lambda_{(1)} - \gamma\mu_{(1)} - \lambda_{(0)})\mu_{(0)}}{2w}\Big) exp\Big(- \frac{(\tilde{y}_{(0)}^2 - 2\tilde{y}_{(0)}\mu_{(0)})}{2\sigma^2|\mu_{(0)}|}\Big)$$

For $\mu_{(t)}$

$$p(\mu_{(t)}|\cdot) \propto p(\mu_{(t)}|\mu_{(t-1)}) \times p(y_{(t)}|\mu_{(t)}, \sigma^2) \times p(\mu_{(t+1)}|\mu_{(t)})$$

$$p(\mu_{(t)}|\mu_{(t-1)}) \propto exp\Big(- \frac{\big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2}{2w}\Big)$$

$$\propto exp\Big(- \frac{(\mu_{(t)}^2 + 2(\gamma\lambda_{(t-1)} - \lambda_{(t)} - \gamma\mu_{(t-1)})\mu_{(t)})}{2w}\Big) \quad (1)$$

$$p(y_{(t)}|\mu_{(t)}, \sigma^2) \propto (|\mu_{(t)}|)^{-1/2} exp\Big(- \frac{(y_{(t)} - \mu_{(t)})^2}{2\sigma^2|\mu_{(t)}|}\Big)$$

$$\propto (|\mu_{(t)}|)^{-1/2} exp\Big(- \frac{y_{(t)}^2 - 2y_{(t)}\mu_{(t)}}{2\sigma^2|\mu_{(t)}|}\Big) \quad (2)$$

$$p(\mu_{(t+1)}|\mu_{(t)}) \propto exp\Big(- \frac{\big(\mu_{(t+1)} - \lambda_{(t+1)} - \gamma\mu_{(t)} + \gamma\lambda_{(t)}\big)^2}{2w}\Big)$$

$$\propto exp\Big(- \frac{\gamma^2\mu_{(t)}^2 - 2\gamma(\mu_{(t+1)} - \lambda_{(t+1)} + \gamma\lambda_{(t)})\mu_{(t)}}{2w}\Big) \quad (3)$$

By multiplying (1), (2) and (3),

$$
\begin{aligned}
&p(\mu_{(t)}|\cdot) \propto p(y_{(t)}|\mu_{(t)},\sigma^2) \times p(\mu_{(t)}|\mu_{(t-1)}) \times p(\mu_{(t+1)}|\mu_{(t)}) \\
&\propto \frac{1}{\sqrt{|\mu_{(t)}|}} exp\Big(- \frac{y_{(t)}^2 - 2y_{(t)}\mu_{(t)}}{2\sigma^2|\mu_{(t)}|} - \frac{\mu_{(t)}^2 + 2(\gamma\lambda_{(t-1)} - \lambda_{(t)} - \gamma\mu_{(t-1)})\mu_{(t)}}{2w} - \frac{\gamma^2\mu_{(t)}^2 - 2\gamma(\mu_{(t+1)} - \lambda_{(t+1)} + \gamma\lambda_{(t)})\mu_{(t)}}{2w}\Big) \\
&= \frac{1}{\sqrt{|\mu_{(t)}|}} exp\Big(- \frac{y_{(t)}^2 - 2y_{(t)}\mu_{(t)}}{2\sigma^2|\mu_{(t)}|} - \frac{(1+\gamma^2)\mu_{(t)}^2 - 2\{\gamma(\mu_{(t-1)} - \lambda_{(t-1)} + \mu_{(t+1)} - \lambda_{(t+1)}) + (1+\gamma^2)\lambda_{(t)}\}\mu_{(t)}}{2w}\Big)
\end{aligned}
$$

The above is an unknown distribution function and we use Metropolis-Hastings algorithm. $N\Big(\lambda_{(t)}, \frac{\tau^2}{1-\gamma^2}\Big)$ for the proposed distribution of the algorithm.

The proposed distribution for $\mu_{(t)}$ is $N\Big(\lambda_{(t)}, \frac{\tau^2}{1-\gamma^2}\Big)$, where $\lambda_{(t)} = \beta_0 + \beta_1 t + \beta_2 t^2$.

For $\mu_{(T)}$ where T is the indicator for the last observation,

$$
\begin{aligned}
p(\mu_{(T)}|\cdot) \propto & p(\mu_{(T)}|\mu_{(T-1)}) \times p(y_{(T)}|\mu_{(T)},\sigma^2) \\
& p(\mu_{(T)}|\mu_{(T-1)}) \propto exp\Big(- \frac{\big(\mu_{(T)} - \lambda_{(T)} - \gamma\mu_{(T-1)} + \gamma\lambda_{(T-1)}\big)^2}{2w}\Big) \\
& \propto exp\Big(- \frac{\mu_{(T)}^2 + 2(\gamma\lambda_{(T-1)} - \lambda_{(T)} - \gamma\mu_{(T-1)})\mu_{(T)}}{2w}\Big) \qquad (1) \\
& p(y_{(T)}|\mu_{(T)},\sigma^2) \propto (|\mu_{(T)}|)^{-1/2} exp\Big(- \frac{(y_{(T)} - \mu_{(T)})^2}{2\sigma^2|\mu_{(T)}|}\Big) \\
& \propto (|\mu_{(T)}|)^{-1/2} exp\Big(- \frac{y_{(T)}^2 - 2y_{(T)}\mu_{(T)}}{2\sigma^2|\mu_{(T)}|}\Big) \qquad (2)
\end{aligned}
$$

By multiplying (1) and (2),

$$
\begin{aligned}
p(\mu_{(T)}|\cdot) \propto & p(\mu_{(T)}|\mu_{(T-1)}) \times p(y_{(T)}|\mu_{(T)},\sigma^2) \\
& \propto (|\mu_{(T)}|)^{-1/2} exp\Big(- \frac{(\mu_{(T)} + C)^2}{2w}\Big) exp\Big(- \frac{y_{(T)}^2 - 2y_{(T)}\mu_{(T)}}{2\sigma^2|\mu_{(T)}|}\Big)
\end{aligned}
$$

The above is an unknown distribution function and Metropolis-Hastings algorithm was used with the same proposed distribution as $\mu_{(t)}$

A2. Poisson response dynamic model.

$$p(\beta_0|\cdot) \propto p(\beta_0) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\beta_0) \propto exp\Big(-\frac{(\beta_0 - m_0)^2}{2v_0}\Big) \propto exp\Big(-\frac{\beta_0^2 - 2m_0\beta_0}{2v_0}\Big) \tag{1}$$

$$\begin{aligned}
&p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
&= \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - (\beta_0 + \beta_1 t + \beta_2 t^2) - \gamma(\mu_{(t-1)} - (\beta_0 + \beta_1(t-1) + \beta_2(t-1)^2))\big)^2}{2w}\Big) \\
&= exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \beta_0 - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2\big)^2}{2w}\Big) \\
&\propto exp\Big(-\frac{T(\gamma-1)^2\beta_0^2 + 2\sum_{t=1}^{T}\Big((\gamma-1)\cdot(\mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \beta_2(t-1)^2)\Big)\cdot\beta_0}{2w}\Big)
\end{aligned} \tag{2}$$

$$p(\mu_{(0)}|\beta_0, \gamma, w) \propto exp\Big(-\frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big) \propto exp\Big(-\frac{(1-\gamma^2)(\beta_0^2 - 2\mu_{(0)}\beta_0)}{2w}\Big) \tag{3}$$

By multiplying (1), (2) and (3),

$$\begin{aligned}
&p(\beta_0|\cdot) \propto p(\beta_0) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
&exp\Big(-\frac{\beta_0^2 - 2m_0\beta_0}{2v_0}\Big) \\
&\times exp\Big(-\frac{T(\gamma-1)^2\beta_0^2 + 2\sum_{t=1}^{T}\Big((\gamma-1)(\mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \beta_2(t-1)^2)\Big)\beta_0}{2w}\Big) \\
&\times exp\Big(-\frac{(1-\gamma^2)(\beta_0^2 - 2\mu_{(0)}\beta_0)}{2w}\Big) \\
&= exp\Big(-\frac{w\beta_0^2 - 2m_0 w\beta_0 + v_0 T(\gamma-1)^2\beta_0^2 + 2v\sum_{t=1}^{T}\Big((\gamma-1)C\beta_0\Big) + v(1-\gamma^2)(\beta_0^2 - 2\mu_{(0)}\beta_0)}{2v_0 w}\Big) \\
&= exp\Big(-\frac{\Big(w + v_0 T(\gamma-1)^2 + v_0(1-\gamma^2)\Big)\beta_0^2 + 2\Big(v_0\sum_{t=1}^{T}\big((\gamma-1)C\big) - v_0(1-\gamma^2)\mu_{(0)} - m_0 w\Big)\beta_0}{2v_0 w}\Big)
\end{aligned}$$

where $C = \mu_{(t)} - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_1(t-1) + \beta_2(t-1)^2$

Now, let $w + v_0 T(\gamma-1)^2 + v_0(1-\gamma^2)$ be A,

$-2\Big(v_0\sum_{t=1}^{T}\big((\gamma-1)C\big) - v_0(1-\gamma^2)\mu_{(0)} - m_0 w\Big)$ be B

and $2v_0 w$ be D.

Then the full conditional posterior distribution for β_0 will be:

$$p(\beta_0|\cdot) \propto exp\Big(-\frac{(\beta_0 + \frac{B}{2A})^2}{2\frac{D}{2A}}\Big)$$

$$\beta_0|\cdot \sim N\Big(-\frac{B}{2A}, \frac{D}{2A}\Big)$$

For β_1,

$$p(\beta_1|\cdot) \propto p(\beta_1) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$p(\beta_1) \propto exp\Big(-\frac{(\beta_1 - m_1)^2}{2v_1}\Big) \propto exp\Big(-\frac{\beta_1^2 - 2m_1\beta_1}{2v_1}\Big) \tag{1}$$

$$p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$= \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \lambda_{(t)} - \gamma(\mu_{(t-1)} - \lambda_{(t-1)})\big)^2}{2w}\Big)$$

$$= exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \beta_0 - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2\big)^2}{2w}\Big)$$

$$= exp\Big(-\frac{\sum_{t=1}^{T}\Big(\big(\gamma(t-1) - t\big)\beta_1 + \big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)\Big)^2}{2w}\Big) \propto$$

$$exp\Big(-\frac{\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)\big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)\beta_1\big)}{2w}\Big) \tag{2}$$

By multiplying (1) and (2),

$$p(\beta_1|\cdot) \propto p(\beta_1) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w)$$

$$exp\Big(-\frac{\beta_1^2 - 2m_1\beta_1}{2v_1}\Big) \times$$

$$exp\Big(-\frac{\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)\big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)\beta_1\big)}{2w}\Big)$$

$$= exp\Big(-\frac{w\beta_1^2 - 2m_1 w\beta_1 + v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2\beta_1^2 + 2v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C\beta_1}{2v_1 w}\Big)$$

$$= exp\Big(-\frac{\big(v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2 + w\big)\beta_1^2 + 2(v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C - m_1 w)\beta_1}{2v_1 w}\Big)$$

where $C = \big(\mu_{(t)} - \beta_0 - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_2(t-1)^2\big)$.

Now, let $v_1\sum_{t=1}^{T}\big(\gamma(t-1) - t\big)^2 + w$ be A',

$2(v_1\sum_{t=1}^{t}\big(\gamma(t-1) - t\big)C - m_1 w)$ be B'

and $2v_1 w$ be D.

Then the full conditional posterior distribution for β_1 will be:

$$p(\beta_1|\cdot) \propto exp\Big(-\frac{(\beta_1 + \frac{B'}{2A'})^2}{2 \cdot \frac{D}{2A'}}\Big)$$

$$\beta_1|\cdot \sim N\Big(-\frac{B'}{2A'}, \frac{D}{2A'}\Big)$$

For β_2,

$$\begin{aligned}
p(\beta_2|\cdot) \propto & p(\beta_2) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& p(\beta_2) \propto exp\Big(-\frac{(\beta_2 - m_2)^2}{2v_2}\Big) \propto exp\Big(-\frac{\beta_2^2 - 2m_2\beta_2}{2v_2}\Big) \qquad (1) \\
& p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& = \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)})\big)^2}{2w}\Big) \\
& = exp\Big(-\frac{\sum_{t=1}^{T}\big(\mu_{(t)} - \beta_0 - \beta_1 t - \beta_2 t^2 - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1) + \gamma\beta_2(t-1)^2\big)^2}{2w}\Big) \\
& = exp\Big(-\frac{\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\beta_2^2 + 2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C)\beta_2}{2w}\Big) \qquad (2)
\end{aligned}$$

By multiplying (1) and (2),

$$\begin{aligned}
& p(\beta_2|\cdot) \propto p(\beta_2) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& exp\Big(-\frac{\beta_2^2 - 2m_2\beta_2}{2v_2}\Big)\times \\
& exp\Big(-\frac{\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\beta_2^2 + 2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C)\beta_2}{2w}\Big) \\
& = exp\Big(-\frac{\Big(w + v_2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2\Big)\beta_2^2 + 2\Big(v_2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)\cdot C) - m_2 w\Big)\beta_2}{2v_2 w}\Big)
\end{aligned}$$

where $C = \big(\mu_{(t)} - \beta_0 - \beta_1 t - \gamma\mu_{(t-1)} + \gamma\beta_0 + \gamma\beta_1(t-1)\big)$

Now, let $w + v_2\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)^2$ be A'',

$2\Big(v_2\cdot\sum_{t=1}^{T}(\gamma(t-1)^2 - t^2)C) - m_2 w\Big)$ be B''

and $2v_2 w$ be D.

Then the full conditional posterior distribution for β_1 will be:

$$\begin{aligned}
p(\beta_2|\cdot) &\propto exp\Big(-\frac{(\beta_2 + \frac{B''}{2A''})^2}{2\cdot\frac{D}{2A''}}\Big) \\
\beta_2|\cdot &\sim N\Big(-\frac{B''}{2A''}, \frac{D}{2A''}\Big)
\end{aligned}$$

For γ,

$$
\begin{aligned}
p(\gamma|\cdot) \propto & p(\gamma) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& p(\gamma) \propto 1 && (1) \\
& p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& = \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) \propto exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)})\big)^2}{2w} \Big) \\
& \propto exp\Big(- \frac{\sum_{t=1}^{T} \big(\gamma(\lambda_{(t-1)} - \mu_{(t-1)}) + (\mu_{(t)} - \lambda_{(t)})\big)^2}{2w} \Big) && (2) \\
& p(\mu_{(0)}|\beta_0, \gamma, w) \propto \sqrt{1-\gamma^2} exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w} \Big) && (3)
\end{aligned}
$$

By multiplying (1), (2) and (3),

$$
\begin{aligned}
& p(\gamma|\cdot) \propto p(\gamma) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& = \sqrt{1-\gamma^2} exp\Big(- \frac{\sum_{t=1}^{T} \big(\gamma(\lambda_{(t-1)} - \mu_{(t-1)}) + (\mu_{(t)} - \lambda_{(t)})\big)^2 + (1-\gamma^2)\beta_0^2}{2w} \Big)
\end{aligned}
$$

The above is an unknown distribution function, so we use meteropolis-hastings algorithm. $Unif(-1, 1)$ was used for the proposed distribution for Metropolis-Hastings algorithm.

For $\tau^2 = w$,

$$
\begin{aligned}
& \tau \sim Unif(0, u) \\
& f_\tau(\tau) = \frac{1}{u} \\
& w = \tau^2 \\
& \tau = w^{\frac{1}{2}} = g^{-1}(w) \\
& f_w(w) = f_\tau(g^{-1}(w)) |\frac{dg^{-1}(w)}{dw}| = \frac{1}{2u} w^{-\frac{1}{2}}
\end{aligned}
$$

Therefore, the full conditional posterior distribution for w is:

$$
\begin{aligned}
p(w|\cdot) \propto & p(w) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w) \\
& p(w) \propto w^{-1/2} && (1) \\
& p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) = \\
& \Pi_{t=1}^{T} p(\mu_{(t)}|\mu_{(t-1)}) \propto ((w)^{-T/2}) exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2}{2w} \Big) && (2) \\
& p(\mu_{(0)}|\beta_0, \gamma, w) \propto (w)^{-1/2} exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w} \Big) && (3)
\end{aligned}
$$

By multiplying (1), (2) and (3),

$$
\begin{aligned}
&p(w|\cdot) \propto p(w) \times p(\mu_{(1)}, \mu_{(2)}, \mu_{(3)}, ..., \mu_{(T)}|\beta_0, \beta_1, \beta_2, \gamma, w) \times p(\mu_{(0)}|\beta_0, \beta_1, \beta_2, \gamma, w)\\
&\big((w)^{-T/2}\big) exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2}{2w}\Big) \times (w)^{-1} exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big)\\
&= \big((w)^{-(T/2)-1}\big) exp\Big(- \frac{\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2 + (1-\gamma^2)(\mu_{(0)} - \beta_0)^2}{2w}\Big)
\end{aligned}
$$

Now, let $\sum_{t=1}^{T} \big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2 + (1-\gamma^2)(\mu_{(0)} - \beta_0)^2$ be A'''.

Then the full conditional posterior distribution for w will be:

$$
\begin{aligned}
&p(w|\cdot) \propto \big((w)^{-(T/2)-1}\big) exp\Big(- \frac{A'''/2}{w}\Big)\\
&w = \tau^2|\cdot \sim inv.gamma\Big(\frac{T}{2}, \frac{A'''}{2}\Big)
\end{aligned}
$$

For $\mu_{(0)}$,

$$
\begin{aligned}
p(\mu_{(0)}|\cdot) \propto & p(\mu_{(0)}) \times p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w)\\
& p(\mu_{(0)}) \propto exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)} - \lambda_{(0)})^2}{2w}\Big) \propto exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)}^2 - 2\lambda_{(0)}\mu_{(0)})}{2w}\Big) \quad (1)\\
& p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w) \propto exp\Big(- \frac{\big(\mu_{(1)} - \lambda_{(1)} - \gamma\mu_{(0)} + \gamma\lambda_{(0)}\big)^2}{2w}\Big)\\
& \propto exp\Big(- \frac{\gamma^2\mu_{(0)}^2 - 2\gamma(\mu_{(1)} - \lambda_{(1)} + \gamma\lambda_{(0)})\mu_{(0)}}{2w}\Big) \quad (2)
\end{aligned}
$$

By multiplying (1) and (2),

$$
\begin{aligned}
&p(\mu_{(0)}|\cdot) \propto p(\mu_{(0)}) \times p(\mu_{(1)}|\mu_{(0)}, \beta_0, \beta_1, \beta_2, \gamma, w)\\
&= exp\Big(- \frac{(1-\gamma^2)(\mu_{(0)}^2 - 2\lambda_{(0)}\mu_{(0)})}{2w}\Big) exp\Big(- \frac{\gamma^2\mu_{(0)}^2 - 2\gamma(\mu_{(1)} - \lambda_{(1)} + \gamma\lambda_{(0)})\mu_{(0)}}{2w}\Big)\\
&\propto exp\Big(- \frac{\mu_{(0)}^2 + 2(\gamma\lambda_{(1)} - \gamma\mu_{(1)} - \lambda_{(0)})\mu_{(0)}}{2w}\Big)
\end{aligned}
$$

Now, let $\gamma\lambda_{(1)} - \gamma\mu_{(1)} - \lambda_{(0)}$ be A''''.

Then the full conditional posterior distribution for $\mu_{(0)}$ will be:

$$
\begin{aligned}
&p(\mu_{(0)}|\cdot) \propto exp\Big(- \frac{\mu_{(0)}^2 + 2A''''\mu_{(0)}}{2w}\Big)\\
&\mu_{(0)}|\cdot \sim N\Big(- A'''', w\Big)
\end{aligned}
$$

For $\mu_{(t)}$

$$
\begin{aligned}
p(\mu_{(t)}|\cdot) \propto & p(\mu_{(t)}|\mu_{(t-1)}) \times p(y_{(t)}|\mu_{(t)}) \times p(\mu_{(t+1)}|\mu_{(t)}) \\
& p(\mu_{(t)}|\mu_{(t-1)}) \propto exp\Big(- \frac{\big(\mu_{(t)} - \lambda_{(t)} - \gamma\mu_{(t-1)} + \gamma\lambda_{(t-1)}\big)^2}{2w}\Big) \\
& \propto exp\Big(- \frac{\big(\mu_{(t)}^2 + 2(\gamma\lambda_{(t-1)} - \lambda_{(t)} - \gamma\mu_{(t-1)})\mu_{(t)}\big)}{2w}\Big) \qquad (1) \\
& p(y_{(t)}|\mu_{(t)}) \propto exp\Big(- exp(\mu_{(t)}) + \mu_{(t)}y_{(t)}\Big) \qquad (2) \\
& p(\mu_{(t+1)}|\mu_{(t)}) \propto exp\Big(- \frac{\big(\mu_{(t+1)} - \lambda_{(t+1)} - \gamma\mu_{(t)} + \gamma\lambda_{(t)}\big)^2}{2w}\Big) \\
& \propto exp\Big(- \frac{\gamma^2\mu_{(t)}^2 - 2\gamma(\mu_{(t+1)} - \lambda_{(t+1)} + \gamma\lambda_{(t)})\mu_{(t)}}{2w}\Big) \qquad (3)
\end{aligned}
$$

By multiplying (1), (2) and (3),

$$
\begin{aligned}
& p(\mu_{(t)}|\cdot) \propto p(\mu_{(t)}|\mu_{(t-1)}) \times p(y_{(t)}|\mu_{(t)}, \sigma^2) \\
& \propto exp\Big(- \frac{\mu_{(t)}^2 + 2(\gamma\lambda_{(t-1)} - \lambda_{(t)} - \gamma\mu_{(t-1)})\mu_{(t)}}{2w} - \frac{\gamma^2\mu_{(t)}^2 - 2\gamma(\mu_{(t+1)} - \lambda_{(t+1)} + \gamma\lambda_{(t)})\mu_{(t)}}{2w} - exp(\mu_{(t)}) + \mu_{(t)}y_{(t)}\Big)
\end{aligned}
$$

The above is an unknown distribution function and $N\Big(\lambda_{(t)}, \frac{w}{1-\gamma^2}\Big)$ was used for Metropolis-Hastings algorithm.

For $\mu_{(T)}$ where T is the indicator for the last observation,

$$
\begin{aligned}
p(\mu_{(T)}|\cdot) \propto & p(\mu_{(T)}|\mu_{(T-1)}) \times p(y_{(T)}|\mu_{(T)}) \\
& p(\mu_{(T)}|\mu_{(T-1)}) \propto exp\Big(- \frac{\big(\mu_{(T)} - \lambda_{(T)} - \gamma\mu_{(T-1)} + \gamma\lambda_{(T-1)}\big)^2}{2w}\Big) \\
& \propto exp\Big(- \frac{\mu_{(T)}^2 + 2(\gamma\lambda_{(T-1)} - \lambda_{(T)} - \gamma\mu_{(T-1)})\mu_{(T)}}{2w}\Big) \qquad (1) \\
& p(y_{(T)}|\mu_{(T)}) \propto exp\Big(- exp(\mu_{(T)}) + \mu_{(T)}y_{(T)}\Big) \qquad (2)
\end{aligned}
$$

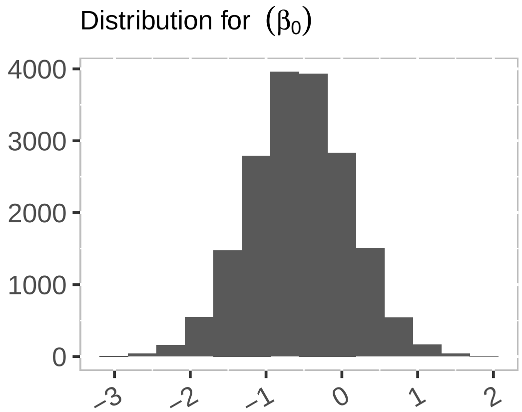
By multiplying (1) and (2),

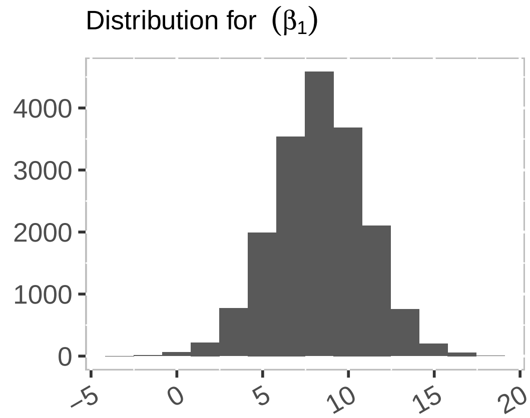
$$
\begin{aligned}
& p(\mu_{(T)}|\cdot) \propto p(\mu_{(T)}|\mu_{(T-1)}) \times p(y_{(T)}|\mu_{(T)}) \\
& \propto exp\Big(- \frac{\mu_{(T)}^2 + 2(\gamma\lambda_{(T-1)} - \lambda_{(T)} - \gamma\mu_{(T-1)})\mu_{(T)}}{2w} - exp(\mu_{(T)}) + \mu_{(T)}y_{(T)}\Big)
\end{aligned}
$$

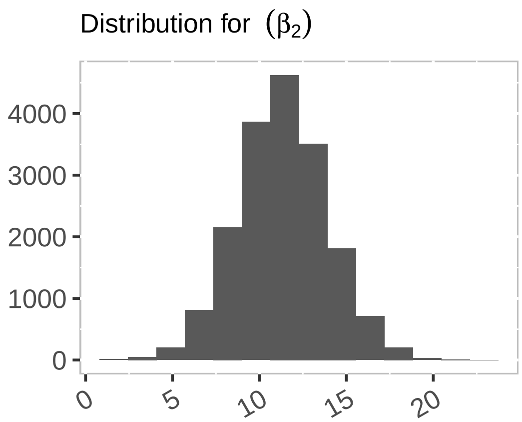
The above is an unknown distribution function and the same proposed distribution as $\mu_{(t)}$ was used.

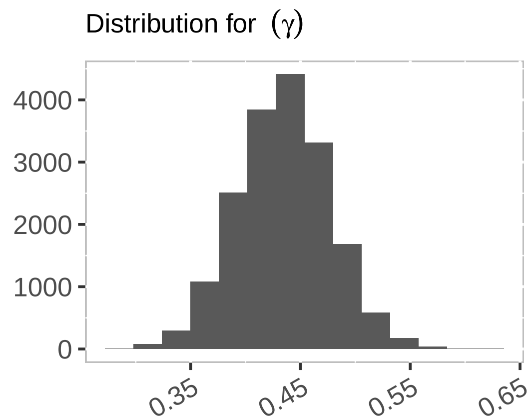
B. Posterior distributions

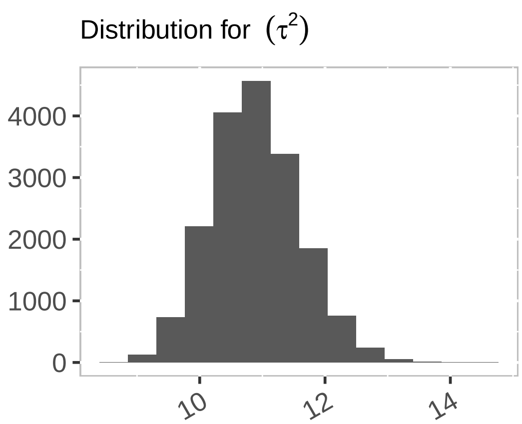
B.1 Posterior distributions for Model 1, Region 1 (Eurasian countries)

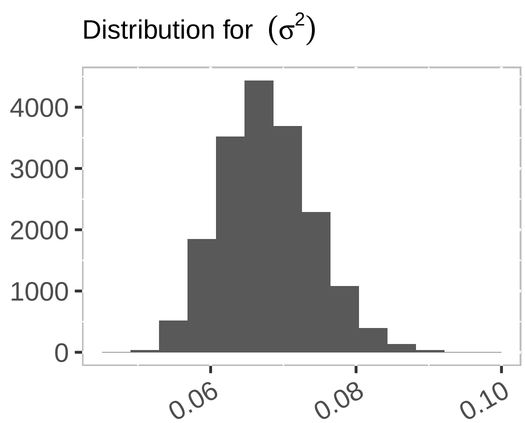




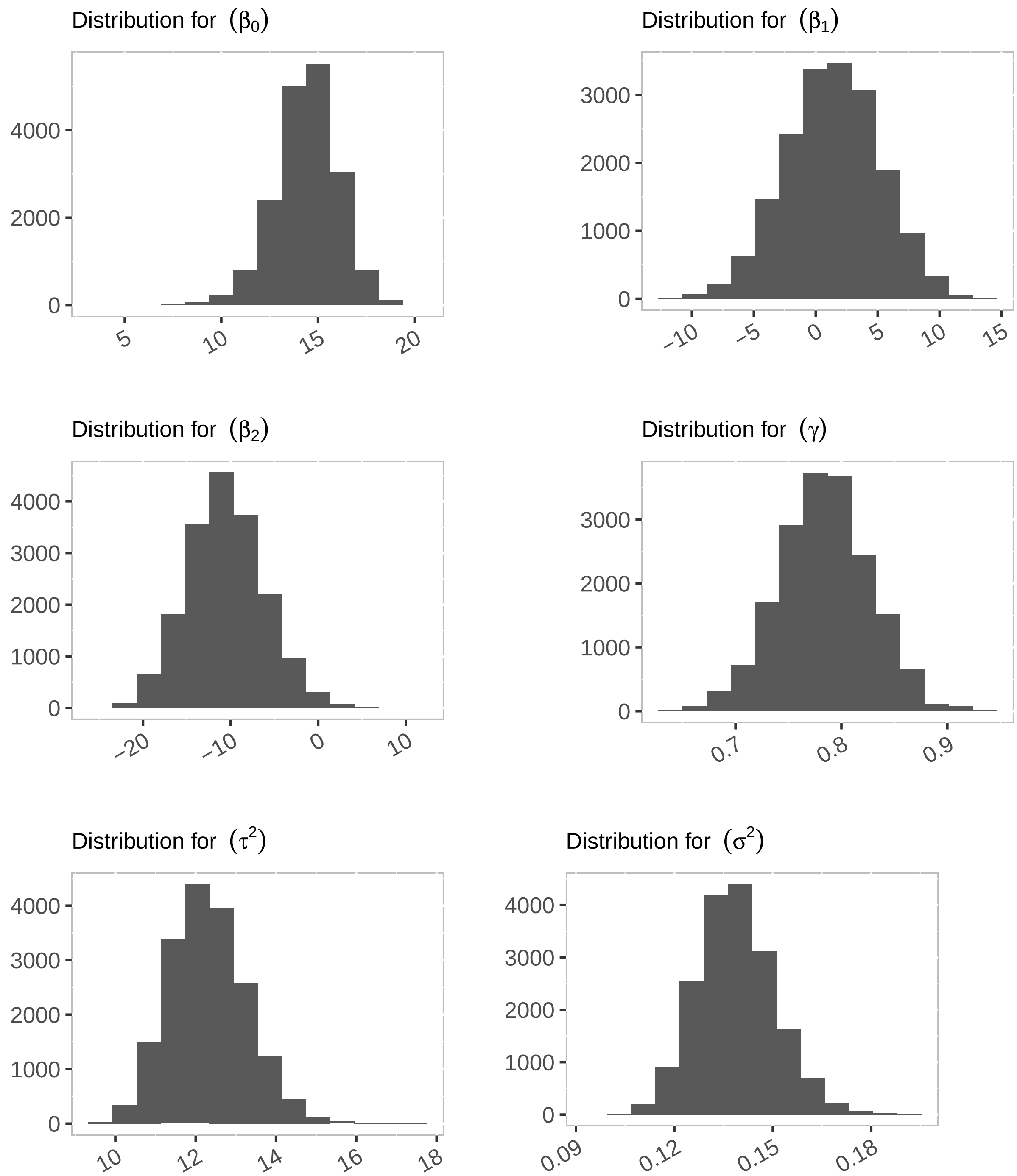




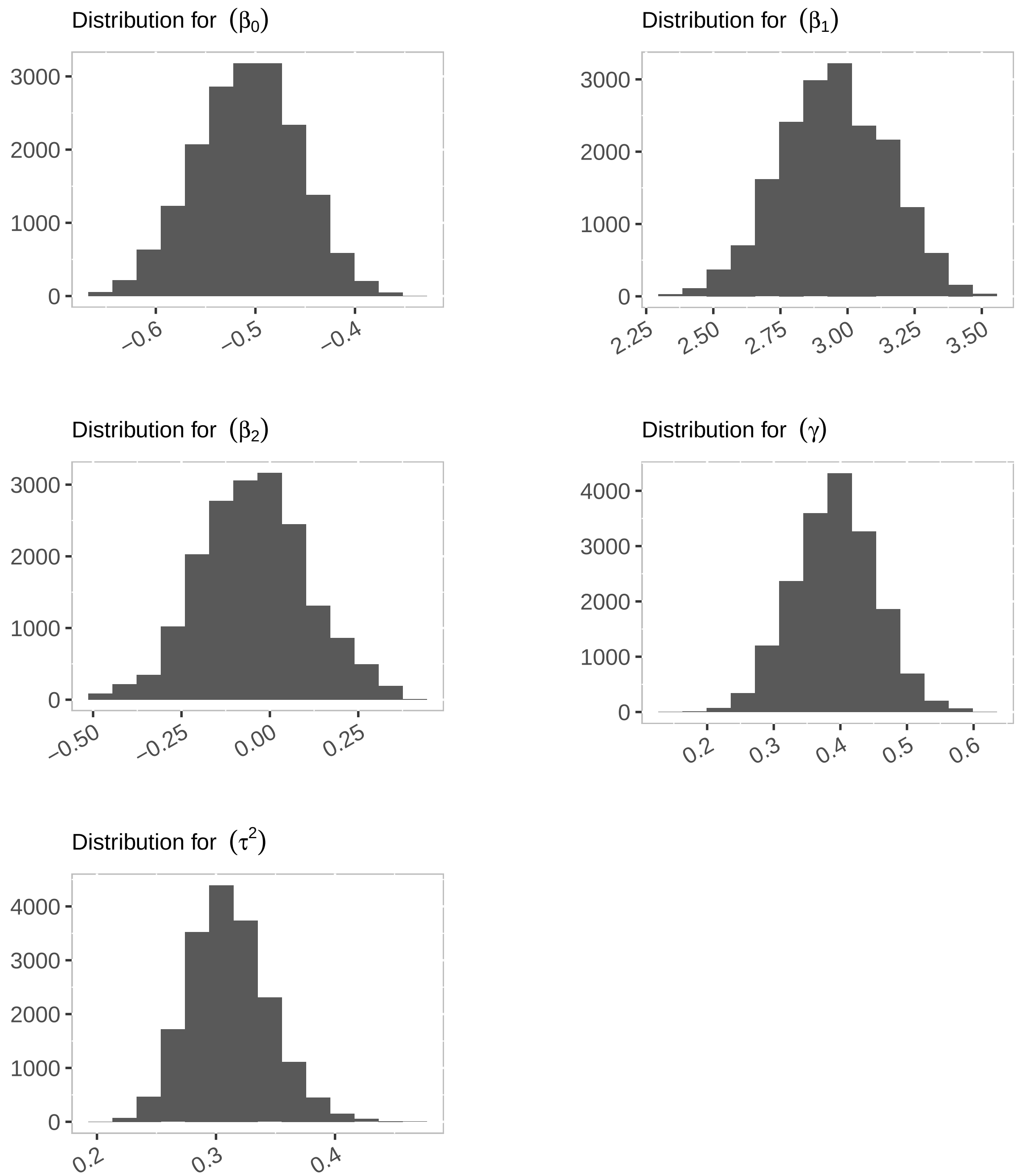




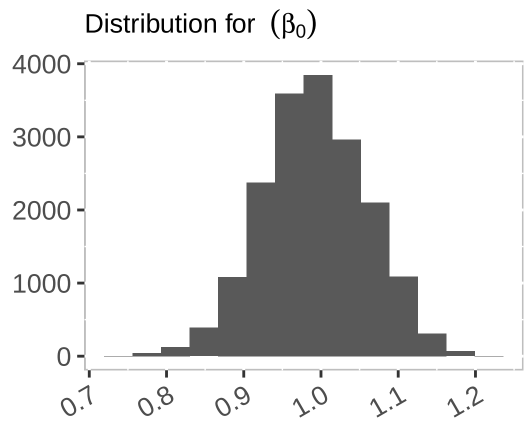
B.2 Posterior distributions for Model 1, Region 2 (Latin American countries)

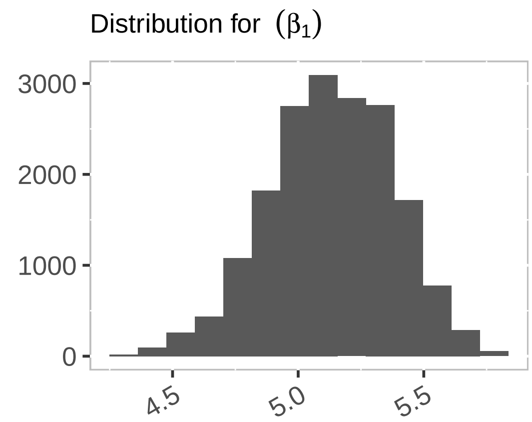


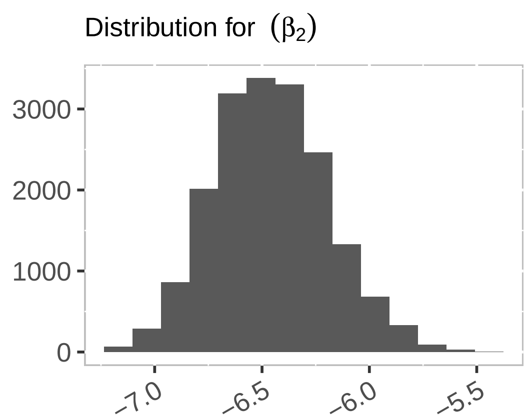
B.3 Posterior distributions for Model 2, Region 1 (Eurasian countries)

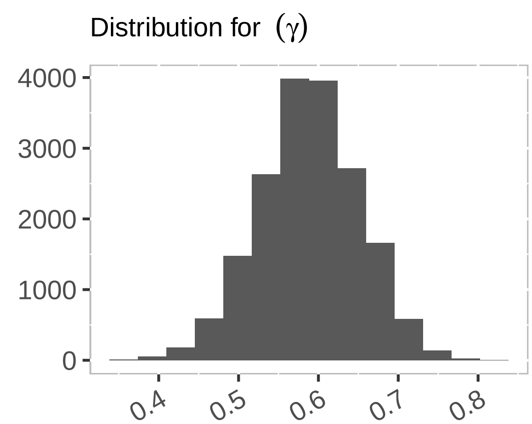


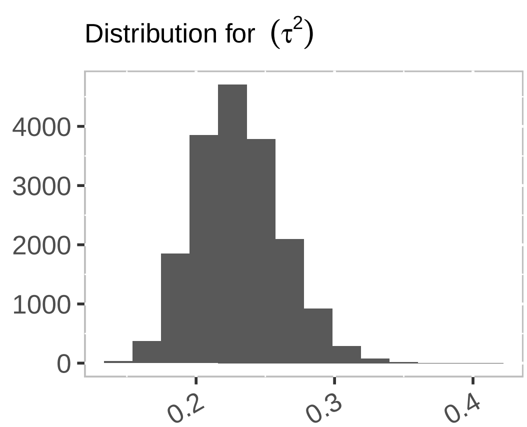
B.4 Posterior distributions for Model 2, Region 2 (Latin American countries)





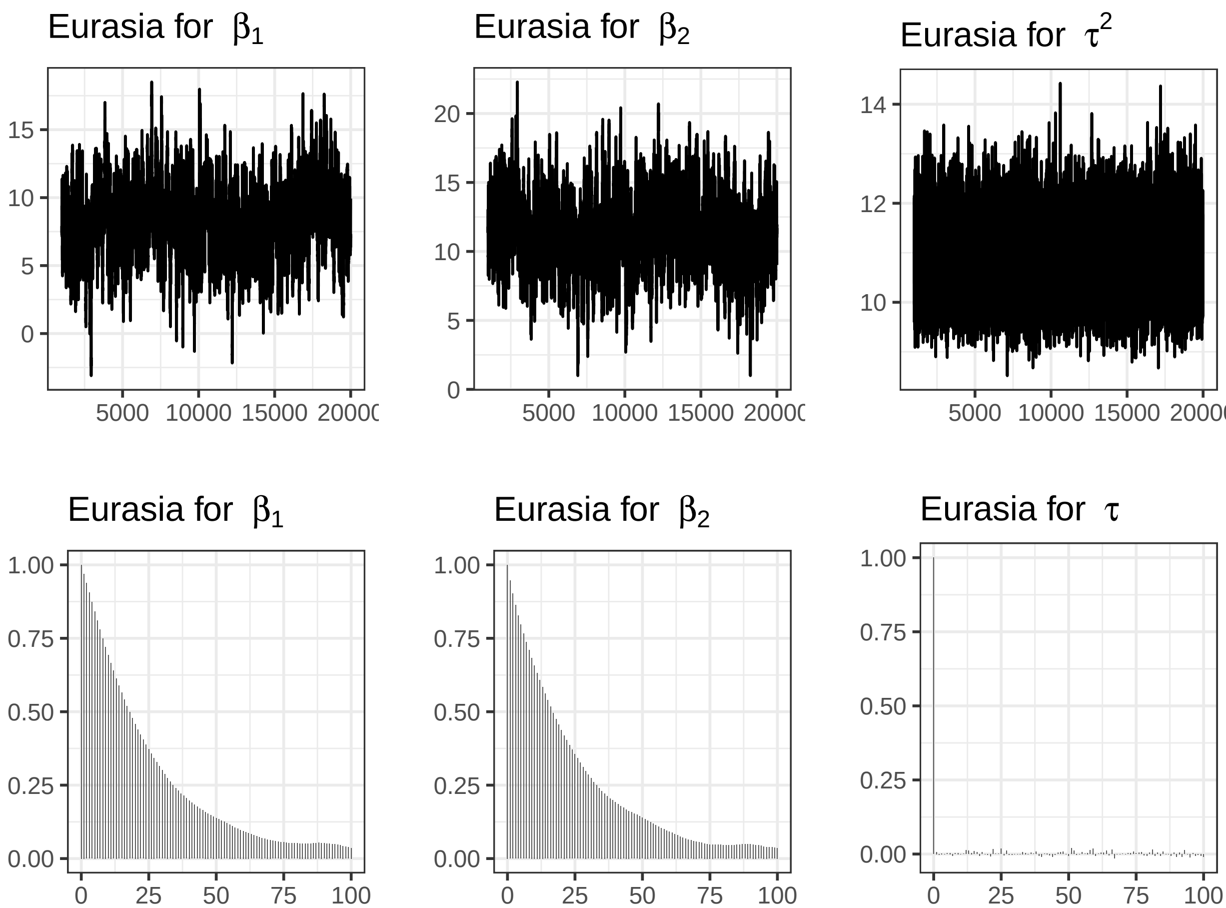




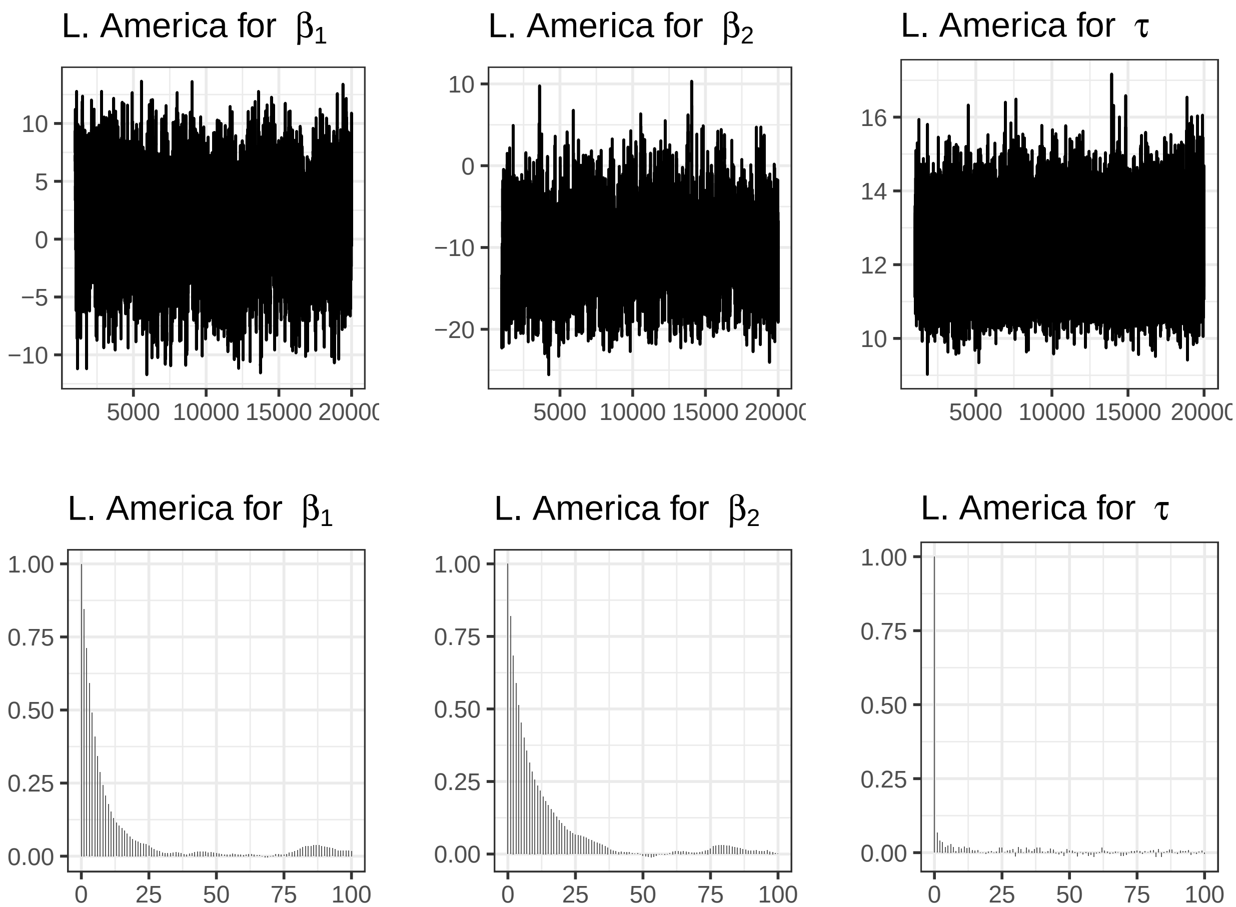


C. Trace Plots and autocorrelation Plots.

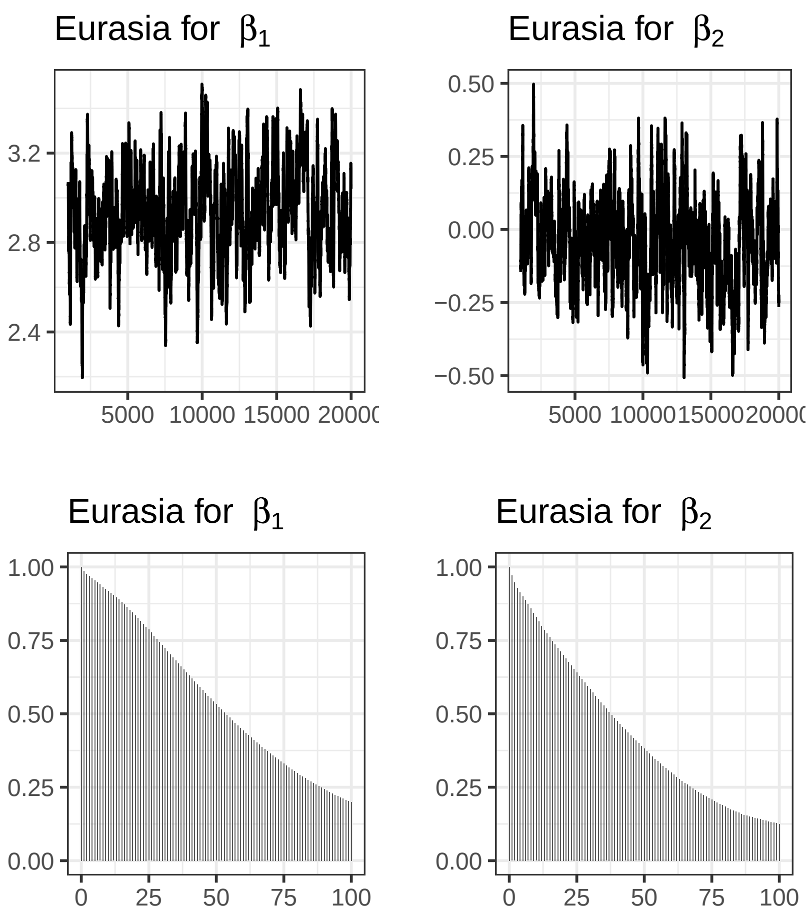
C.1 For Gaussian latent variable model - Eurasian countries.



C.2 For Gaussian latent variable model - Latin American countries.



C.3 For Poisson response dynamic model - Eurasian countries.



C.4 For Poisson response dynamic model - Latin American countries.